

# GLOBAL JOURNAL

OF COMPUTER SCIENCE AND TECHNOLOGY: C

---

## Software & Data Engineering

Instrumental Decoding

Path Guiding System

**Highlights**

Support Vector Machine

Supervised Learning

Discovering Thoughts, Inventing Future

VOLUME 14

ISSUE 1

VERSION 1.0



GLOBAL JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY: C  
SOFTWARE & DATA ENGINEERING

---



GLOBAL JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY: C  
SOFTWARE & DATA ENGINEERING

---

VOLUME 14 ISSUE 1 (VER. 1.0)

OPEN ASSOCIATION OF RESEARCH SOCIETY

© Global Journal of Computer Science and Technology. 2014 .

All rights reserved.

This is a special issue published in version 1.0 of "Global Journal of Computer Science and Technology" By Global Journals Inc.

All articles are open access articles distributed under "Global Journal of Computer Science and Technology"

Reading License, which permits restricted use. Entire contents are copyright by of "Global Journal of Computer Science and Technology" unless otherwise noted on specific articles.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without written permission.

The opinions and statements made in this book are those of the authors concerned. Ultraculture has not verified and neither confirms nor denies any of the foregoing and no warranty or fitness is implied.

Engage with the contents herein at your own risk.

The use of this journal, and the terms and conditions for our providing information, is governed by our Disclaimer, Terms and Conditions and Privacy Policy given on our website <http://globaljournals.us/terms-and-condition/menu-id-1463/>

By referring / using / reading / any type of association / referencing this journal, this signifies and you acknowledge that you have read them and that you accept and will be bound by the terms thereof.

All information, journals, this journal, activities undertaken, materials, services and our website, terms and conditions, privacy policy, and this journal is subject to change anytime without any prior notice.

Incorporation No.: 0423089  
License No.: 42125/022010/1186  
Registration No.: 430374  
Import-Export Code: 1109007027  
Employer Identification Number (EIN):  
USA Tax ID: 98-0673427

## Global Journals Inc.

(A Delaware USA Incorporation with "Good Standing"; Reg. Number: 0423089)

Sponsors: *Open Association of Research Society*  
*Open Scientific Standards*

### *Publisher's Headquarters office*

**Global Journals Headquarters**  
301st Edgewater Place Suite, 100 Edgewater Dr.-Pl,  
**Wakefield MASSACHUSETTS, Pin: 01880,**  
**United States of America**

*USA Toll Free: +001-888-839-7392*  
*USA Toll Free Fax: +001-888-839-7392*

### *Offset Typesetting*

**Global Journals Incorporated**  
2nd, Lansdowne, Lansdowne Rd., Croydon-Surrey,  
**Pin: CR9 2ER, United Kingdom**

### *Packaging & Continental Dispatching*

**Global Journals**  
E-3130 Sudama Nagar, Near Gopur Square,  
**Indore, M.P., Pin:452009, India**

### *Find a correspondence nodal officer near you*

To find nodal officer of your country, please email us at [local@globaljournals.org](mailto:local@globaljournals.org)

### *eContacts*

Press Inquiries: [press@globaljournals.org](mailto:press@globaljournals.org)  
Investor Inquiries: [investers@globaljournals.org](mailto:investers@globaljournals.org)  
Technical Support: [technology@globaljournals.org](mailto:technology@globaljournals.org)  
Media & Releases: [media@globaljournals.org](mailto:media@globaljournals.org)

### *Pricing (Including by Air Parcel Charges):*

*For Authors:*

22 USD (B/W) & 50 USD (Color)  
*Yearly Subscription (Personal & Institutional):*  
200 USD (B/W) & 250 USD (Color)

INTEGRATED EDITORIAL BOARD  
(COMPUTER SCIENCE, ENGINEERING, MEDICAL, MANAGEMENT, NATURAL  
SCIENCE, SOCIAL SCIENCE)

**John A. Hamilton, "Drew" Jr.,**  
Ph.D., Professor, Management  
Computer Science and Software  
Engineering  
Director, Information Assurance  
Laboratory  
Auburn University

**Dr. Henry Hexmoor**  
IEEE senior member since 2004  
Ph.D. Computer Science, University at  
Buffalo  
Department of Computer Science  
Southern Illinois University at Carbondale

**Dr. Osman Balci, Professor**  
Department of Computer Science  
Virginia Tech, Virginia University  
Ph.D. and M.S. Syracuse University,  
Syracuse, New York  
M.S. and B.S. Bogazici University,  
Istanbul, Turkey

**Yogita Bajpai**  
M.Sc. (Computer Science), FICCT  
U.S.A. Email:  
yogita@computerresearch.org

**Dr. T. David A. Forbes**  
Associate Professor and Range  
Nutritionist  
Ph.D. Edinburgh University - Animal  
Nutrition  
M.S. Aberdeen University - Animal  
Nutrition  
B.A. University of Dublin- Zoology

**Dr. Wenying Feng**  
Professor, Department of Computing &  
Information Systems  
Department of Mathematics  
Trent University, Peterborough,  
ON Canada K9J 7B8

**Dr. Thomas Wischgoll**  
Computer Science and Engineering,  
Wright State University, Dayton, Ohio  
B.S., M.S., Ph.D.  
(University of Kaiserslautern)

**Dr. Abdurrahman Arslanyilmaz**  
Computer Science & Information Systems  
Department  
Youngstown State University  
Ph.D., Texas A&M University  
University of Missouri, Columbia  
Gazi University, Turkey

**Dr. Xiaohong He**  
Professor of International Business  
University of Quinnipiac  
BS, Jilin Institute of Technology; MA, MS,  
PhD., (University of Texas-Dallas)

**Burcin Becerik-Gerber**  
University of Southern California  
Ph.D. in Civil Engineering  
DDes from Harvard University  
M.S. from University of California, Berkeley  
& Istanbul University

**Dr. Bart Lambrecht**

Director of Research in Accounting and Finance  
Professor of Finance  
Lancaster University Management School  
BA (Antwerp); MPhil, MA, PhD  
(Cambridge)

**Dr. Carlos García Pont**

Associate Professor of Marketing  
IESE Business School, University of Navarra  
Doctor of Philosophy (Management),  
Massachusetts Institute of Technology (MIT)  
Master in Business Administration, IESE,  
University of Navarra  
Degree in Industrial Engineering,  
Universitat Politècnica de Catalunya

**Dr. Fotini Labropulu**

Mathematics - Luther College  
University of Regina  
Ph.D., M.Sc. in Mathematics  
B.A. (Honors) in Mathematics  
University of Windsor

**Dr. Lynn Lim**

Reader in Business and Marketing  
Roehampton University, London  
BCom, PGDip, MBA (Distinction), PhD,  
FHEA

**Dr. Mihaly Mezei**

ASSOCIATE PROFESSOR  
Department of Structural and Chemical  
Biology, Mount Sinai School of Medical  
Center  
Ph.D., Eötvös Loránd University  
Postdoctoral Training,  
New York University

**Dr. Söhnke M. Bartram**

Department of Accounting and Finance  
Lancaster University Management School  
Ph.D. (WHU Koblenz)  
MBA/BBA (University of Saarbrücken)

**Dr. Miguel Angel Ariño**

Professor of Decision Sciences  
IESE Business School  
Barcelona, Spain (Universidad de Navarra)  
CEIBS (China Europe International Business School).  
Beijing, Shanghai and Shenzhen  
Ph.D. in Mathematics  
University of Barcelona  
BA in Mathematics (Licenciatura)  
University of Barcelona

**Philip G. Moscoso**

Technology and Operations Management  
IESE Business School, University of Navarra  
Ph.D in Industrial Engineering and Management, ETH Zurich  
M.Sc. in Chemical Engineering, ETH Zurich

**Dr. Sanjay Dixit, M.D.**

Director, EP Laboratories, Philadelphia VA  
Medical Center  
Cardiovascular Medicine - Cardiac  
Arrhythmia  
Univ of Penn School of Medicine

**Dr. Han-Xiang Deng**

MD., Ph.D  
Associate Professor and Research  
Department Division of Neuromuscular  
Medicine  
Department of Neurology and Clinical  
Neuroscience  
Northwestern University  
Feinberg School of Medicine

**Dr. Pina C. Sanelli**

Associate Professor of Public Health  
Weill Cornell Medical College  
Associate Attending Radiologist  
NewYork-Presbyterian Hospital  
MRI, MRA, CT, and CTA  
Neuroradiology and Diagnostic  
Radiology  
M.D., State University of New York at  
Buffalo, School of Medicine and  
Biomedical Sciences

**Dr. Roberto Sanchez**

Associate Professor  
Department of Structural and Chemical  
Biology  
Mount Sinai School of Medicine  
Ph.D., The Rockefeller University

**Dr. Wen-Yih Sun**

Professor of Earth and Atmospheric  
SciencesPurdue University Director  
National Center for Typhoon and  
Flooding Research, Taiwan  
University Chair Professor  
Department of Atmospheric Sciences,  
National Central University, Chung-Li,  
TaiwanUniversity Chair Professor  
Institute of Environmental Engineering,  
National Chiao Tung University, Hsin-  
chu, Taiwan.Ph.D., MS The University of  
Chicago, Geophysical Sciences  
BS National Taiwan University,  
Atmospheric Sciences  
Associate Professor of Radiology

**Dr. Michael R. Rudnick**

M.D., FACP  
Associate Professor of Medicine  
Chief, Renal Electrolyte and  
Hypertension Division (PMC)  
Penn Medicine, University of  
Pennsylvania  
Presbyterian Medical Center,  
Philadelphia  
Nephrology and Internal Medicine  
Certified by the American Board of  
Internal Medicine

**Dr. Bassey Benjamin Esu**

B.Sc. Marketing; MBA Marketing; Ph.D  
Marketing  
Lecturer, Department of Marketing,  
University of Calabar  
Tourism Consultant, Cross River State  
Tourism Development Department  
Co-ordinator , Sustainable Tourism  
Initiative, Calabar, Nigeria

**Dr. Aziz M. Barbar, Ph.D.**

IEEE Senior Member  
Chairperson, Department of Computer  
Science  
AUST - American University of Science &  
Technology  
Alfred Naccash Avenue – Ashrafieh

## PRESIDENT EDITOR (HON.)

---

### **Dr. George Perry, (Neuroscientist)**

Dean and Professor, College of Sciences

Denham Harman Research Award (American Aging Association)

ISI Highly Cited Researcher, Iberoamerican Molecular Biology Organization

AAAS Fellow, Correspondent Member of Spanish Royal Academy of Sciences

University of Texas at San Antonio

Postdoctoral Fellow (Department of Cell Biology)

Baylor College of Medicine

Houston, Texas, United States

## CHIEF AUTHOR (HON.)

---

### **Dr. R.K. Dixit**

M.Sc., Ph.D., FICCT

Chief Author, India

Email: [authorind@computerresearch.org](mailto:authorind@computerresearch.org)

## DEAN & EDITOR-IN-CHIEF (HON.)

---

### **Vivek Dubey(HON.)**

MS (Industrial Engineering),

MS (Mechanical Engineering)

University of Wisconsin, FICCT

Editor-in-Chief, USA

[editorusa@computerresearch.org](mailto:editorusa@computerresearch.org)

### **Sangita Dixit**

M.Sc., FICCT

Dean & Chancellor (Asia Pacific)

[deanind@computerresearch.org](mailto:deanind@computerresearch.org)

### **Suyash Dixit**

(B.E., Computer Science Engineering), FICCTT

President, Web Administration and

Development , CEO at IOSRD

COO at GAOR & OSS

### **Er. Suyog Dixit**

(M. Tech), BE (HONS. in CSE), FICCT

SAP Certified Consultant

CEO at IOSRD, GAOR & OSS

Technical Dean, Global Journals Inc. (US)

Website: [www.suyogdixit.com](http://www.suyogdixit.com)

Email: [suyog@suyogdixit.com](mailto:suyog@suyogdixit.com)

### **Pritesh Rajvaidya**

(MS) Computer Science Department

California State University

BE (Computer Science), FICCT

Technical Dean, USA

Email: [pritesh@computerresearch.org](mailto:pritesh@computerresearch.org)

### **Luis Galárraga**

J!Research Project Leader

Saarbrücken, Germany



## CONTENTS OF THE VOLUME

---

- i. Copyright Notice
- ii. Editorial Board Members
- iii. Chief Author and Dean
- iv. Table of Contents
- v. From the Chief Editor's Desk
- vi. Research and Review Papers
  1. Student's Learning Progression through Instrumental Decoding of Mathematical Ideas. *1-42*
  2. A Hybrid Random Forest based Support Vector Machine Classification Supplemented by Boosting. *43-53*
  3. Location based Path Guiding System for the Visually Impaired People. *55-59*
  4. A Framework for Context-Aware Semi Supervised Learning. *61-70*
  5. Supervised Classification of Remote Sensed Data using Support Vector Machine. *71-76*
- vii. Auxiliary Memberships
- viii. Process of Submission of Research Paper
- ix. Preferred Author Guidelines
- x. Index



# Student's Learning Progression Through Instrumental Decoding of Mathematical Ideas

By Stavroula Patsiomitou

*University of Ioannina, Greece*

**Abstract** - The current study aims to focus on mathematical tasks for students' mathematical literacy and problem solving literacy. Excerpts are presented from dynamic hypothetical learning paths [DHLP] s and students' learning progression. The excerpts center around activities aimed to develop the students' geometrical thinking through the development of their ability to solve real-world problems. The students cooperated in class or worked individually to represent the images using their static or dynamic means and tools (e.g. compass and ruler, a computing environment, interactive boards, dynamic geometry software). My further aim was the students to utilize transformation processes for representations by instrumentally decoding their ideas on static and dynamic objects.

**Keywords** : *linking visual active representations, learning progression, 'dynamic' hypothetical learning path, teaching cycle.*

**GJCST-C Classification** : *K.8.0*



*Strictly as per the compliance and regulations of:*



# Student's Learning Progression through Instrumental Decoding of Mathematical Ideas

Stavroula Patsiomitou

**Abstract-** The current study aims to focus on mathematical tasks for students' mathematical literacy and problem solving literacy. Excerpts are presented from dynamic hypothetical learning paths [DHLP] s and students' learning progression. The excerpts center around activities aimed to develop the students' geometrical thinking through the development of their ability to solve real-world problems. The students cooperated in class or worked individually to represent the images using their static or dynamic means and tools (e.g. compass and ruler, a computing environment, interactive boards, dynamic geometry software). My further aim was the students to utilize transformation processes for representations by instrumentally decoding their ideas on static and dynamic objects. An important role for the students' cognitive development was the design of propositions and theorems (e.g the Pythagorean Theorem), through Linking Visual Active Representations (LVAR). Especially for the latter option an essential role has played the dynamic geometry software, Geometer's Sketchpad. Furthermore, the paper provides examples that contain rich mathematical material; therefore, student's mathematical modeling through instrumental decoding of mathematical ideas is the means of reinforcing students' conceptual knowledge.

**Keywords:** *linking visual active representations, learning progression, 'dynamic' hypothetical learning path, teaching cycle.*

## I. INTRODUCTION

The current study aims to focus on mathematical tasks for students' development of geometrical thinking "in the process of developing and refining a learning progression to build a coherent [geometry] curriculum [connected with the other areas of mathematics] and the associated instructional materials" (Krajcik, Shin, Stevens & Short, 2009, p.27). For this, the paper describes excerpts from predicted [hypothetical] learning paths (/trajectories) "through which the learning might proceed. [The learning trajectories are hypothetical as it was not] knowable in advance" (Simon, 1995, p.135). Furthermore these learning paths are dynamic, as instructional DG (Dynamic Geometry) -- as The Geometer's Sketchpad (Jackiw, 1991) -- activities are incorporated. Therefore, they could be defined as Dynamic Hypothetical Learning Paths (DHLPs). I have initially been designed and modified the paths as a result of interactions with the students that participated, adding the destinations

that were not known in advance to me (Simon, 1995, p.137).

The learning paths "are subsets of [a] learning progression [...] as it requires developing and testing an entire series of learning [paths] that describe specifically how to move students toward conceptual understanding of the big idea[s] in [mathematics and particularly in geometry]" (Krajcik, Shin, Stevens & Short, 2009, p.27).

Furthermore, Simon (ibid.) developed the idea of a teaching cycle and created a diagram in order to represent the way that a learning trajectory is an ongoing modification of three components: "(a) the learning goal that defines the direction, (b) the learning activities and (c) the hypothetical learning process"(Simon, 1995, p. 136). Mathematics tasks are related to the teacher's mathematical and pedagogical knowledge. According to Simon (1995) "the ingredient necessary in order to initiate mathematics learning is *pedagogy*" (p. 115, italics in original manuscript). Furthermore, teacher's knowledge about effective mathematical pedagogy influences their instructional practices (e.g., Simon & Shifter, 1991; Carpenter, Fennema, Peterson, Chiang, & Loef, 1989).

The DHLPs incorporated real-world problems or simulations of problems in the DGS environment that had been analyzed and designed in terms of (a) the students' van Hiele (vH) levels of thinking, starting from the lower vH levels to elicit higher vH levels, (b) their sequential conceptual content, and (c) the student's comprehension of the links between representations and mathematical meanings conceptually and procedurally.

Points of departure for the anticipation of the DHLPs were the questions:

- Do students understand the mathematical components of modeling when they see real-world environments' images [-representations]?
- What mental activities will the students develop when they participate in these learning [/ instructional] activities?
- What mathematical representations are most appropriate for student learning?
- How important is the role of a dynamic geometry program to reorganize students' mental representations?
- Does students' actual learning correspond with what was anticipated?

*Author:* University of Ioannina, Ioannina, Greece.  
*e-mail:* spatsiomitou@gmail.com

- How effective is the teaching and learning process using linking visual active representations (LVARs) (e.g., Patsiomitou, 2008 a, b, c, d, 2009, 2010a, b, 2011, 2012a, b, 2013; Patsiomitou & Koleza, 2008, 2009; Patsiomitou & Emvalotis, 2009 a, b, c, d, Patsiomitou & Emvalotis 2010 a, b; Patsiomitou, Barkatsas, & Emvalotis, 2010) to overcome cognitive obstacles and develop understanding of the mathematical concepts?
- Which mathematical competencies are developed through the DHLHP?
- What is the need for students' learning and understanding in upper-class curriculum processes for innovative learning when new practices and ideas (i.e., fractal activities) are incorporated?

The paper provides a link between "learning with actions" and the implementation of mathematics in educational and pedagogical contexts, answering the question of if we could incorporate real world mathematics into everyday school practices.

Mathematics is part of every day children's lives. The mathematics is obvious to them or not, sometimes is not perceived as they are implicit. On the other hand

*"School courses and books have presented 'mathematics' as a series of apparently meaningless technical procedures. [...] Just as a phrase loses meaning or acquires an unintended meaning when removed from its context, so mathematics detached from its rich intellectual setting in the culture of our civilization and reduced to a series of techniques has been grossly distorted"* (Kline, 1990, p.15-16).

Moreover, the fractal approach that is presented here reflects Kaput's (1992) writing on the importance of technology in mathematics education, concerning the feasibility of innovative practices emanating from technological advents, which were otherwise impracticable. Ferrara, Pratt, & Robutti (2006) also, suggest that "what promotes change is the curricular project in which technology is inserted, and in particular, the didactic sequences planned by the teachers in order to introduce [...] concepts, which use technology as a support." (p. 258).

The article does not intend to present the extended results of the research process but rather the theoretical perspective that underpins the teaching cycle (Simon, 1995) and the role of dynamic LVAR in students' cognitive development.

In the next sections, the article will begin with an articulation of the constructivist perspective that underpins the student learning process and my decision for the selection of activities. A review will be provided of mathematical competencies and the role of modeling processes in the DGS environment with the utilization of LVARs.

## II. THEORETICAL UNDERPINNING

### a) Student's cognitive development

During the past several decades, researchers were concerned about the difficulties their students have faced when attempting geometry problems (e.g., Hoffer, 1981; Usiskin, 1982). This consistent result comes about through students' difficulty releasing their thoughts from a concrete frame (White & Mitchelmore, 2010, p. 206), and failure to develop the deductive reasoning (Peirce, 1998/1903) required. This prevents them from engaging in the abstract process (e.g., Skemp, 1986; White & Mitchelmore, 2010) that is required for the study of the conceptual structure of geometry.

According to Piaget (1937/1971), students' cognitive development depends on their biological maturity. That students' cognitive development depends on the teaching process was argued by Dina van Hiele-Geldof and Pierre van Hiele in their dissertations in 1957 (Fuys, Geddes & Tischler, 1988). Dina van Hiele-Geldof (Fuys, Geddes & Tischler, 1984) in her dissertation had the objective to investigate the improvement of learning performance by a change in the learning method. Central to this model, is the description of the five levels of thought development which are: Level 1 (recognition or visualization), Level 2 (analysis), Level 3 (ordering), Level 4 (deduction) and Level 5 (rigor).

Battista uses "constructivist constructs such as levels of abstraction to describe students' progression through the van Hiele levels" (Battista, 2011, p.515). He "has elaborated the original van Hiele levels to carefully trace students' progress in moving from informal intuitive conceptualizations of 2D geometric shapes to the formal property-based conceptual system used by mathematicians" (Battista, 2007, p.851).

He separated each phase in subphases (Battista, 2007). I briefly report Battista's first three levels elaboration, which are the most pertinent to secondary students, below:

*Level 1* (Visual-Holistic Reasoning) is separated into sublevel 1.1. (prerecognition) and sublevel 1.2 (recognition). (p.851).

*Level 2* (Analytic-Componential Reasoning) is separated into sublevel 2.1 (Visual-informal componential reasoning), sublevel 2.2 (Informal and insufficient-formal componential reasoning) sublevel 2.3 (Sufficient formal property-based reasoning). According to Battista (2007) "Students [acquire through instruction] a) an increasing ability and inclination to account for the spatial structure of shapes by analyzing their parts and how their parts are related and b) an increasing ability to understand and apply formal geometric concepts in analyzing relationships between parts of shapes". (pp.851-852).

*Level 3* (Relational –Inferential Property-Based Reasoning) into sublevel 3.1 (Empirical relations), sublevel 3.2 (Componential analysis), sublevel 3.3

(Logical inference) and sublevel 3.4 (Hierarchical shape, classification based on logical inference). According to Battista (2007) "Students explicitly interrelate and make inferences about geometric properties of shapes. [...] The verbally-stated properties themselves are interiorized so that they can be meaningfully decomposed, analyzed, and applied to various shapes". (pp. 852-853).

Researchers have shown that students "often fail in the construction of a geometric configuration which is essential for the solution of the underlying geometric problem" (Schumann & Green, 1994, p.204). This happens because students at the lower levels "identify, describe, and reason about shapes and other geometric configurations according to their appearance as visual wholes" (Battista, 2007, p.851). According to van Hiele (1986) "when after some time, the concepts are sufficiently clear, pupils can begin to describe them. With this the properties possessed by the geometric figures that have been dealt with are successively mentioned and so become explicit. The figure becomes the representative of all these properties: It gets what we call the "symbol character". In this stage the comprehension of the figure means the knowledge of all these properties as a unity. [...] When the symbol character of many geometric figures have become sufficiently clear to the pupils, the possibility is born that they also get a signal character". This means that the symbols can be anticipated. [...] When this orientation has been sufficiently developed, when the figures sufficiently act as signals, then, for the first time geometry can be practiced as a logical topic" (p. 168).

Many researchers (e.g., Guitierrez & Jaime, 1998; Govender & De Villiers, 2002, 2004; Patsiomitou, 2008, 2012a, b, 2013; Patsiomitou & Emvalotis, 2010 a, b) describe student's processes of constructing definitions and justification at every van Hiele level as they develop geometrical thought. This evolution of students' formulation of definitions, justification, and reasoning was adopted by this study as the characteristic that would indicate their movement through several van Hiele levels. For definitions, I adopted Govender and De Villiers' (2004) clarification (see Patsiomitou, 2013). In addition, dynamic perceptual definition (e.g. Patsiomitou, 2013, p.806) is the term for the process by which the student informally 'defines' a geometrical object by using the tools of the software.

#### b) *The development of student's mathematical competencies*

Another point of view suggests that the development of student's geometrical thinking results from the development of their competencies in mathematical thinking and reasoning, argumentation, modeling etc. Therefore, if the teaching process of students is aimed to develop their competencies, then it leads to the development of their geometrical thinking.

Many researchers (e.g, Burkhardt, 1981; Pierce & Stacey, 2009) have highlighted the idea of solving problems in the real world as essential to understanding and learning mathematics, as well as "a key ability for citizens [who are prepared to make] judgments and decisions" (Stacey, 2012, p.3).

According to De Corte, Verschaffel & Greer (2000), the implementation of the mathematics to solve real world problems can be useful "as a complex process involving a number of phases: understanding the situation described; constructing a mathematical model that describes the essence of those elements and relations embedded in the situation that are relevant; working through the mathematical model to identify what follows from it; interpreting the outcome of the computational work to arrive at a solution to the practical situation that gave rise to the mathematical model; evaluating that interpreted outcome in relation to the original situation; and communicating the interpreted results".(p.1).

Through the solution of the real world problems, students will be assessed regarding their competency for horizontal and vertical mathematization (Jupri, Drijvers, & van den Heuvel-Panhuizen, 2012). "The difficulty in horizontal mathematization concerns students' difficulty in going from the world of real phenomena to the world of symbols and vice versa. The difficulty in vertical mathematization concerns students' difficulty in dealing with the process of moving within the symbolic world (Treffers, 1987; Van den Heuvel-Panhuizen, 2003)" (Reported in <http://igitur-archive.library.uu.nl/math/2013-0304-200631/12102012.pdf>).

As I previously mentioned, my further aims, were the student's mathematical literacy and problem-solving literacy. The latter PISA (Programme for International Student Assessment) definition of mathematical literacy is as follows (OECD, 2010):

*"Mathematical literacy is an individual's capacity to formulate, employ, and interpret mathematics in a variety of contexts. It includes reasoning mathematically and using mathematical concepts, procedures, facts, and tools to describe, explain, and predict phenomena. It assists individuals to recognise the role that mathematics plays in the world and to make the well-founded judgments and decisions needed by constructive, engaged and reflective citizens." (p. 4)*

It is very important for the students to develop their modeling competency in order to transform real-world problems from the three-dimensional world to the two-dimensional world of the paper and pencil [or DG] environment. Additionally, it is important for them to be able to process in an abstract way.

Epigrammatically, the students, through the problems that will be presented below, will be assessed with regard to the development of the following

competencies (OECD, 2006), that have been analyzed from Niss (1999) and his colleagues but similar formulations can be found in the work of many others (e.g., Neubrand et al. 2001):

Briefly, these competencies can be described as an individual student's ability to (e.g., Niss, 1999, 2003; Neubrand et al. 2001):

*Mathematical thinking and reasoning:*... mastering mathematical modes of thought; posing questions characteristic of mathematics; knowing the kind of answers that mathematics offers, distinguishing among different kinds of statements; understanding and handling the extent and limits of mathematical concepts; generalizing results to larger classes of objects.

*Mathematical reasoning and argumentation:* ...knowing what proofs are; knowing how proofs differ from other forms of mathematical reasoning; following and assessing chains of arguments; having a feel for heuristics; creating and expressing mathematical arguments; devising formal and informal mathematical arguments, and transforming heuristic arguments to valid proofs, i.e. proving statements.

*Mathematical communication:* ... being able to communicate, in, with, and about mathematics; expressing oneself in a variety of ways in oral, written, and other visual form; understanding someone else's work.

*Modelling competency:*... being able to analyse and build mathematical models concerning other subjects or practice areas; structuring the field to be modeled; translating reality into mathematical structures; interpreting mathematical models in terms of context or reality; working with models; validating models; reflecting, analyzing, and offering critiques of models or solutions; reflecting on the modeling process; communicating about the model and its results; monitoring and controlling the entire modeling process.

*Problem posing and handling competency:*... problem identifying, posing, specifying; solving different kinds of mathematical problems.

*Representation competency:*... being able to handle different representations of mathematical entities; decoding, encoding, translating, distinguishing between, and interpreting different forms of representations of mathematical objects and situations as well as understanding the relationship among different representations; choosing and switching between representations.

*Symbol and formalism competency:*... decoding and interpreting symbolic and formal mathematical language, and understanding its relations to natural language; understanding the nature and rules of formal mathematical systems (both syntax and semantics); translating from natural language to formal/symbolic language; handling and manipulating statements and expressions containing symbols and formulae.

*Communicating in, with, and about mathematics competency:* .... understanding others' written, visual or oral 'texts', in a variety of linguistic registers, about matters having a mathematical content; expressing oneself, at different levels of theoretical and technical precision, in oral, visual or written form, about such matters.

*Aids and tools competency:*...being able to make use of and relate to the aids and tools of mathematics, including technology when appropriate.

The visualization competency and the competency of students to develop recursive processes conceptually and structurally is [also] very important for the solution of problems with fractal constructions.

In a paper of PME conference (Patsiomitou, 2011) I had also distinguished the kinds of apprehension when selecting software objects. Competence in the DGS environment depends on the competence of the cognitive analysis which students bring to bear when decoding the utilization of software tools, based on Duval's (1995) semiotic analysis of students' apprehension of a geometric figure. During the development of a construction in a DGS environment, I believe that the student has to develop three kinds of apprehension when selecting software objects which accord with the types of cognitive apprehension outlined by Duval (1995, pp.145-147) namely perceptual, sequential, discursive, and operative apprehension. In concrete terms, the competence of instrumental decoding in the software's constructions depends on (Patsiomitou, 2011, p.363):

a) the sequential apprehension of the tools selection (i.e. s/he has to select point C and segment AB and then the command (fig. 1) meaning that s/he has to follow a predetermined order); b) the verbal apprehension of the tools selection which means the student has to verbalize this process, (i.e. s/he says "I am going to select point C and the segment AB") and c) a place way type of elements operation on the figure (i.e. when s/he transforms the orientation of the elements to apply the command selecting point B and the opposite side AC, for example in fig.4) due to his/her perceptual apprehension (fig. 2, 4). Then s/he has constructed the operative apprehension of the figure's elements for the construction, meaning the competence to operate the construction.

c) *Is learning and knowledge development a cognitive process? The role of teacher in learning process*

Van Hiele theory has its roots in constructivist theories. Cognitive constructivism is connected with the work of Piaget's (1937/1971) and his views as 'constructivist'. Bruner's (1961, 1966) proposal of discovery learning [as 'constructionist'] is based on prior knowledge and the understanding of a concept, which [through discovery] grows and deepens. According to Bruner (1986) "learning is a social process

in which children grow into the intellectual life of those around them" (Clements & Battista, 1990, p.6).

The sociocultural approach has its roots in Vygotsky (1987) who focuses on the acquisition of mathematical understanding as a product of social interactions. Von Glasersfeld (1995) a radical constructivist is differentiated from the work of Piaget as he argues that "knowledge [does not represent an independent world, instead] represents something that [...] we can do in our experiential world" (p.6).

Building on the concepts mentioned above, the concept of social constructivism is a complex process, while being interactive, constructivist and sociocultural (e.g., Yackel, Cobb, Wood, Wheatley & Merkel 1990; Cobb, Yackel & Wood 1992; Yackel, Rasmussen & King 2001; Yackel & Rasmussen 2002; Jaworski, 2003). According to sociocultural and interactive approaches, learning is a part of the culture (Steffe & Gale, 1995) in which the students construct knowledge through their participation in social practices (e.g. social class environment) (Cobb & Bauersfeld, 1995, p.4). "A social-constructivist perspective sees discussion, negotiation and argumentation in inquiry and investigation practices to underpin knowledge growth in mathematics, in teaching mathematics and in mathematics teacher education" (e.g., Cobb & Bowers, 1999; Lampert, 1998; Wood, 1999 in Jaworski, 2003, p. 17).

Besides, learning is an individual constructive process while knowledge is actively constructed by the student; it depends on the individual's personal work and negotiation of mathematical ideas (e.g., Jaworski, 2003). From the perspective of constructivist theories the process of mathematical knowledge and understanding arises as students try to solve math problems during the classroom (Cobb, Yackel, & Wood, 1992; Simon & Shifter, 1991) and is instigated when students confront problematic situations. Knowing therefore is not taken passively by students but in an active way. Learning thus is characterized in Bauersfeld's interactionism view "by the subjective reconstruction of societal means and models through negotiation of meaning in social intervention" (Bauersfeld, 1992, p.39).

Vygotsky (1987) argues that "the child begins to perceive the world not only through his eyes [visually] but also through speech" (p. 32). According to Vygotsky (1987), learning is a complex interplay between scientific and spontaneous use of language.

For this, learning is an internalization of social relations and understanding is a result of common negotiation of concepts created by students while interacting with other students in the class (or group) during the mathematical discussions developed (Bartolini Bussi, 1996).

For the current study, I used the strategy of "thinking aloud" (Hayes & Flower, 1980; Smith & Wedman, 1988) in class or group discussions because I

strongly believe "that this [action] influence[s] students' own use of language". [...]. "Language is important for cognitive development and learning; without it, an individual lacks [an] efficient system for storing certain types of information that are needed for thinking, reasoning, and concept development"(Westwood, 2004, p.141).

Sfard also defines "learning as the process of changing one's discursive ways in a certain well-defined manner" (Sfard, 2001, p.3). According to Sfard (2001) "thinking is a special case of the activity of communicating" [...]"A person who thinks can be seen as communicating with himself/herself, [...] whether the thinking is in words, in images or other form of symbols, [...] as our thinking is [an interactive] dialogical endeavour [through which] we argue..." (p.3); with his/her participation the student in a mathematical discussion s/he "learns to think mathematically" (Sfard, *ibid.*, p. 4). Under this approach, the development of thought occurs through dialogue that develops the subject within himself/herself internally (intrapersonally) or in a group in which s/he participates. Moreover, learning is expanding the capacity for dialectical skills and solving problems that could not previously be solved. Furthermore "putting communication in the heart of mathematics education is likely to change not only the way we teach but also the way we think about learning and about what is being learned" (Sfard, 2001, p.1). Consequently, learning is first and foremost the modification / transformation of the ways we think and how we exchange this thought. Moreover, learning is the capacity of dialectical skills and of problem-solving that could not be solved before.

Goos and her colleagues carried out a series of studies --based on sociocultural perspective-- to investigate the teacher's role, the students' discussion in small groups and the use of technology as a tool that mediates teaching and learning interactions (e.g. Goos, 2004, Goos, Galbraith, Renshaw, & Geiger, 2003). If we take the role of teacher seriously as concerns the realisation and planning of activities then, every activity should be based on geometry exactly as Goldenberg (1999) purports it to be --a fundamental principle. The current study leads us, as Goldenberg (*ibid.*), writes "to select idea-editors that have supported the connections. Tools like Geometer's Sketchpad present geometric structures in an environment that emphasizes the continuous nature of Euclidean space, and thus serve as an excellent bridge between geometry and [the other field of mathematics, as well as] analysis." This is very important for the teaching practice because the construction of the meaning can not only be depended or is located in the tool per se, nor uniquely pinpointed in the interaction of student and tool, but it lies in the schemes of use (e.g., Trouche, 2004) of the tool itself.

Simon (1995) has developed a view of the teacher's role that includes both the psychological and

the social aspects. He supports that "a teacher is directed by his conceptual goals for his students, goals that are constantly being modified" (p.135). I adopted Simon's view for my role as teacher-researcher for the current study. In the next sections it will be articulated the "rational for choosing the particular instructional design; thus I make my design decisions based on my guess of how learning might proceed" (Simon, 1995, p.135).

*d) What are Learning Progressions? What is 'The Teaching Cycle'?*

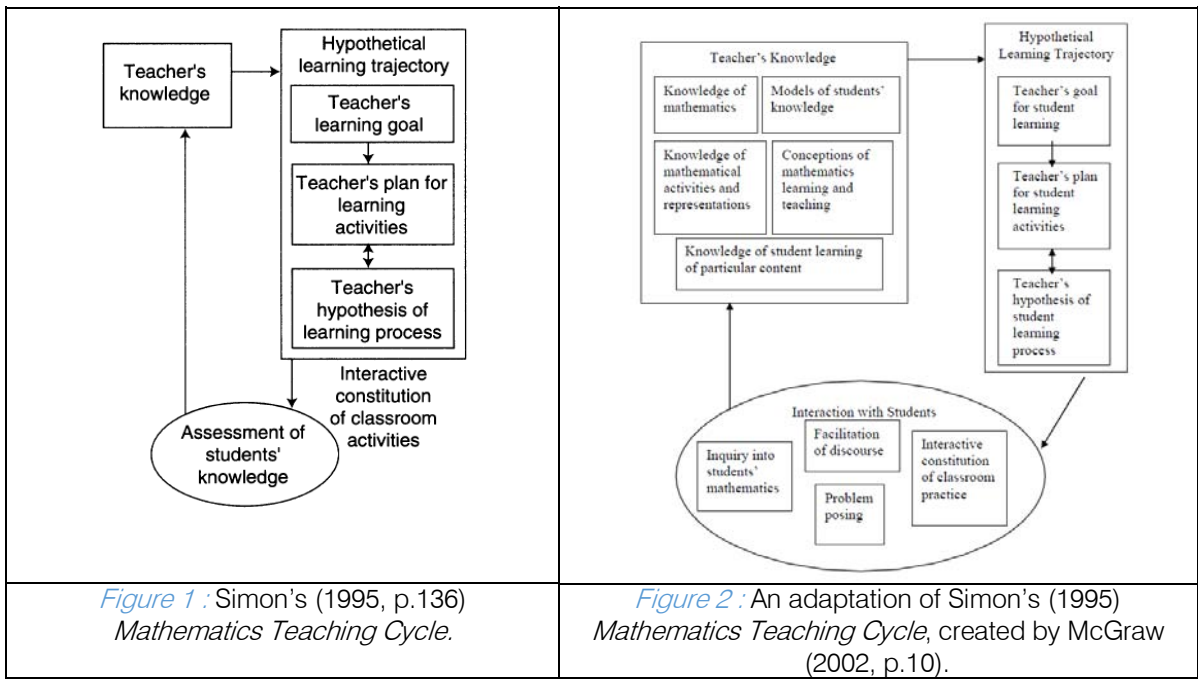
Duschl, Schweingruber, & Shouse (2007) define learning progressions as "descriptions of the successively more sophisticated ways of thinking about a topic that can follow one another as children learn about and investigate a topic over a broad span of time" (Duschl, Schweingruber, & Shouse, 2007, p.214). A learning progression is also committed to the "notion of learning as an ongoing developmental progression. It is designed to help children continually build on, and revise their knowledge and abilities, starting from the initial conceptions about how the world works and curiosity about what they see around them" (National Research Council (NRC), 2010, p.2). Learning progressions has among others students' assessment as important component that aids to "measure student understanding of the key concepts or practices and can track their developmental progress over time" (Corcoran, Mocher & Rogat, 2009, p.15).

The learning progression in mathematics (or other disciplines, e.g language, science) has been built upon the concept of the Assessment Triangle (Pellegrino, Chudowsky & Glaser, 2001) which "explicate three key elements underlying any assessment: (1) a model of student cognition and learning in the domain, (2) a set of beliefs about the kinds of observations that provide evidence of students competencies and (3) an interpretation process for making sense of evidence" (National Research Council, 2001, p.44). As Smith, Wiser, Anderson, & Krajcik (2006) argue the progress through learning progressions depends on instruction as well as the theory of van Hiele supports.

Krajcik, Shin, Stevens & Short, (2009, p.28), have created an illustration of the difference between learning progressions and learning trajectories. As they support "sets of learning trajectories of instructional sequences that describe specific ways of supporting student's learning [constitute] a learning progression" (p.28). Besides Battista (2011) supports that a learning progression differs from a learning trajectory because it has not been designed "to test a curriculum, based on a fixed sequence of learning tasks in that curriculum. [Instead] it is focusing on a formative assessment system that applies to many curricula [...] based on many assessment tasks, not those in a fixed sequence" (p. 513).

In the current study, the vH [learning] progression describes the development of students' thinking through a vH instructional path, which was kept for a year and then repeated the next year. It focuses on describing the evolution of the students' mathematical learning, cognitive structures, and reasoning and has been "developed from examining how students' ideas develop by the analysis of various assessment tasks" (Stevens, Shin & Krajcik, 2009, p.2). I was the activities' instructor as well as student's teacher. The activities are "linked to [students'] curriculum and are being tested by collecting their work from classroom in which the curriculum is being used to specify the expected achievement levels; [the current] study tend to be longitudinal, with changes in student thinking and ability tracked over the course of [...] a year." (Corcoran, Mocher & Rogat, 2009, p.33). According to Neubrand (2004) mathematical achievement can be structured by three "types of mathematical activities," employing only the techniques of modeling and problem-solving activities that use mathematical tools and procedures, call for connections, and utilize mathematical conceptions.





Simon (1995) created a diagram (Figure 1) in order to represent the way that a learning trajectory is an ongoing modification of three components: (a) the learning goal, (b) the learning activities and (c) the hypothetical learning process. McGraw (2002, p.10) in her Ph.D. thesis created an adaptation on Simon's (1995) teaching cycle that is represented in Figure 2, aiming to include the actual discussions with students that "occurred within the 'interaction with students', which influenced the "teacher's knowledge".

What has not been examined is the use of technology in the teaching cycle which plays an important role in the development of discussions, as well as students' vH level.

*e) The modeling process in a DGS environment –What are LVARs?*

From a representational view of learning mathematics the DHLP is supported theoretically by the concept of representation. According to Vergnaud (1987) "representation is an important element in the theory of teaching and learning of mathematics [... especially since they] play an important role in understanding the real world. Representations are provided to the students in different forms (Gagatsis & Spirou, 2000) (e.g., real-world situations, images or diagrams, oral or written symbols).

Vergnaud (1998) claims that "[n] either Piaget nor Vygotsky realized how much cognitive development depends on situations and on the specific conceptualizations that are required to deal with them" (p. 181). Piaget focused on the subjectivity of representation and Vygotsky on a social process of gaining control over external "sign" forms. Children have difficulty to perceive the signs of the meanings in the images of the real world. They perceive them as a whole

image especially at the lower van Hiele levels. When students move to upper van Hiele levels they increase their ability to transform the visual image or drawing (Parsysz, 1988) they perceive, into a figure with concrete properties.

For most researchers, representations can help students to reorganize and translate their ideas using symbols. They are also useful as communication tools (Kaput, 1991) and can function as tools for understanding of concepts, since they help with the communication of ideas and provide a social environment for the development of mathematical discussion. The knowledge of supporting instruments, which are external representational systems for planning activities, allows us to choose between technological tools. The [external] representations facilitate the provision of information about the problem, capture the structure of the problem, and support visual reasoning. On the other hand, the external representations (e.g., formulations or figures) that students construct serve as an indicator of their internal representations, constituting their level of understanding and the developmental level of their geometric thinking.

The use of a computing environment as dynamic geometry facilitates the teaching and learning of Euclidean geometry and helps students overcome the difficulties in translation between representations through automatic translation or "dynamically linking"(Ainsworth, 1999, p. 133) since "visually encode causal, functional, structural, and semantic properties and relationships of a represented world – either abstract or concrete (Glasgow, Narayanan & Chandrasekran, 1995; Peterson, 1996; Card, MacKinlay & Shneiderman, 1999; Cheng, 2002)" (Sedig & Sumner, 2006, p. 2).

Dynamic geometry software has been used broadly in research regarding the teaching and learning process of geometry over the past several decades (see for example the articles written in *Educational Studies in Mathematics* and *International Journal of Computers for Mathematical Learning*) (Leung & Or, 2007, p. 177). Such research with dynamic geometry has verified that the software is useful in provoking cognitive conflicts (e.g., Hadas, Hershkowitz, & Schwarz, 2000; Giraldo, Belfort & Carvalho, 2004), developing students' deductive reasoning (e.g., Hollebrands & Smith, 2009), and developing students' geometrical thinking (e.g., Yousef, 1997; Almeqdadi, 2000; Sinclair, 2001; Patsiomitou, 2008a, 2010a, 2012a,b, 2013), according to the theory of van Hiele. Üstün & Ubuş (2004) consider that "the Geometer's Sketchpad is an important vehicle of technological change in geometry classroom. [...] The shapes are first created and then they are explored, manipulated and transformed to ideal concept". Olkun, Sinoplu & Deryakulu (2005) argued that "the Geometer's Sketchpad is a suitable dynamic environment in which students can explore geometry according to their van Hiele levels" (p.3). The diagrams that are provided to the students in a DGS environment are important spatiovisual representations that facilitate understanding of the problem's information as well as the conceptualization of the problem's structure. In other words the 'dynamic' diagrams support visual reasoning, which aids translation from visual to verbal representations and the construction of meaning.

In Geometer's Sketchpad v4 DGS environment, LVAR (e.g., Patsiomitou, 2008a, b,c,d, 2009, 2010a, b, 2011, 2012a,b, and 2013) are interpreted as a [real world's or not] problem modeling process. The definition of LVAR (Patsiomitou, 2012, p. 76) is given below:

Linking Visual Active Representations are the successive building steps in a dynamic representation of a problem, the steps that are repeated in different problems or steps reversing a procedure in the same phase or between different phases of a hypothetical

learning path. LVARs reveal an increasing structural complexity by conceptually and structurally linking the transformational steps taken by the user (teacher or student) as a result of the interaction techniques provided by the software to externalize the transformational steps s/he has visualized mentally (or exist in his/her mind) or organized as a result of his/her development of thinking and understanding of geometrical concepts.

Real world images (or digital images) "are potential representations [...] and offer the heuristic part of learning" as they "denote something" (Kadunz & Straesser, 2004, p.241, 242). What is important is how the students perceive these potential representations of the environment (natural images or digital), how they use and communicate with each other and how they manage their mental mathematical structures in order to represent the objects. Mogeta, Olivero & Jones (1999) in their report "Providing the Motivation to Prove in a Dynamic Geometry Environment" argue that "setting problem solving within these environments requires a careful design of activities, which need to take into account the interaction between three elements: the dynamic software, as an instance of the milieu, a problem, and a situation, through which the devolution of the problem takes place (Brousseau, 1986)". Most importantly, the diagrams that the students are obliged to translate and the relations that link the objects in the diagram will provide researchers and teachers insights to see their abilities and their weaknesses with respect to the mathematical knowledge that they have structured as a result of the teaching process in class. For this, the verification of students' mistakes and cognitive obstacles during the construction of diagrams will lead us to the reinforcement of the teaching of mathematics in the context of real-world problems.

Vergnaud proposed an approach for investigation in mathematics education, which includes the steps presented in the Figure 3 (Vergnaud, 1988, p.149 reported in Long, 2011, p.123).

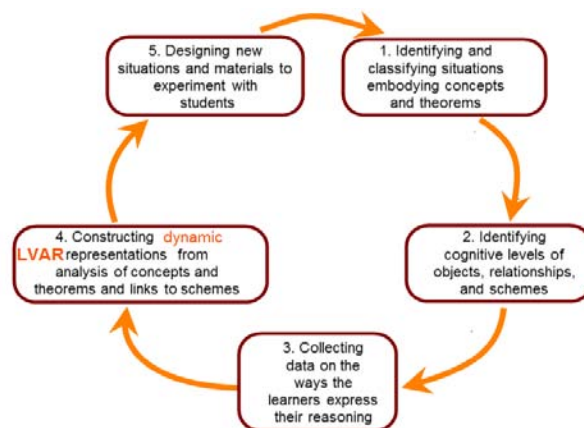


Figure 3 : An adaptation for the current study on Vergnaud's (1988, p.149) approach.

Gonzalez & Herbst (2009) have defined the dynamic diagram as “a diagram made with DGS and that has the potential to be changed in some way by dragging one or more of its parts” (p.154).

According to Mariotti (2000, p.36) “the dragging test, externally oriented at first, is aimed at testing perceptually the correctness of the drawing; as soon as it becomes part of interpersonal activities [...] it changes its function and becomes a sign referring to a meaning, the meaning of the theoretical correctness of the figure.” Hollebrands (2007) also supported that the students in her study “used reactive or proactive strategies when dragging, either in response to or in anticipation of the effects on dragging” (cited in Gonzalez and Herbst, 2009, p.158-159). Building on Mariotti’s considerations and Hollebrands distinction about dragging strategies, in a previous study (Patsiomitou, 2011) I introduced the notions of theoretical dragging (i.e., the student aims to transform a drawing into a figure on screen, meaning s/he intentionally transforms a drawing to acquire additional properties) and experimental dragging (i.e., the student investigates whether the figure (or drawing) has certain properties or whether the modification of the drawing in the picture plane through dragging leads to the construction of another figure or drawing).

Students execute on screen constructions using software’s tools and primitive geometrical objects in an effort to decode their mental representations into software actions. This sense of how the student’s competence at instrumental decoding affects the development of their ability in constructing meanings, may lead to an understanding of how the tools the students use, play a fundamental role as a non linguistic warrant. The construction of a figure on screen in a DGS

environment is a result of a complex process on the student’s part. The student has first to transform the verbal or written formulation (“construct a parallelogram” for example) into a mental image, which is to say an internal representation recalling a prototype image (e.g., Hershkovitz, 1990) that s/he has shaped from a textbook or other authority, before transforming it into an external representation, namely an on-screen construction. This process requires the student to decode their actions using software primitives, functions etc. In order to accomplish a construction in the software the student must acquire the competence for instrumental decoding meaning the competence to transform his/her mental images to actions in the software, using the software’s interaction techniques. Furthermore, dynamic reinvention of knowledge (Patsiomitou, 2012b, p. 57) is the kind of knowledge the students could reinvent by interacting with the artefacts made in a DGS environment, “knowledge for which they themselves are responsible” (Gravemeijer & Terwel, 2000, p.786).

For the representation of student’s argumentation I used a pseudo-Toulmin’s model (Patsiomitou, 2012b, p.57) --based on Toulmin’s model (1958)-- in which: (1) the data could be an element or an object of the dynamic diagram, and (2) a warrant could be a tool or a command that guarantees the result which is the claim (or the resulted formulation). So, students perceive the properties of the rotated segment and during instrumental genesis. According to Schumann (2004) the diagram below “presents an outline of methods and ways of working with DGS in the context of geometry teaching in lower and middle secondary schools; modeling in DGS is supported by all other methods and options”

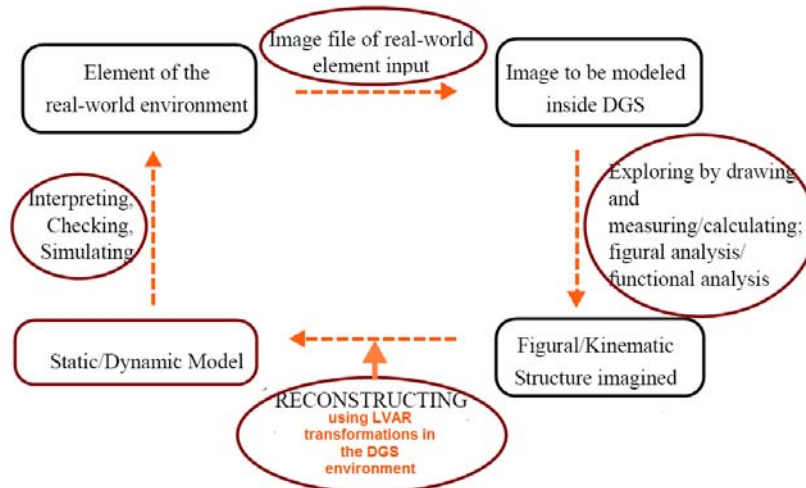


Figure 4 : An adaptation for the current study of Schuman’s (2004, p. 7) modeling process using LVAR

In the figure 4, the factor “reconstructing using LVAR transformations” is the warrant (W) for the claim (static /dynamic model). This means that LVAR transformations guarantee the interpretation of the

dynamic model to students’ mind into external [verbal or iconic] representations.

In the next section a description will be presented of the DHLPs. As it has been told previously

the DHLPs are empirically tested parts of the learning progression. The students applied their mathematical knowledge (pre-existing or not) to solve real-world problems. The modeled problems have been presented in a dynamic geometry environment or the students had to manipulate the images of the real world in their minds in order "to bridge connections between the pure world of mathematics --with fixed solutions and "perfect" forms-- and the more messy, ambiguous, or subjective world of experience" (Sinclair & Jackiw, 2007). In every situation, the experience with a real or simulated object played a major role for the construction of students' knowledge.

### III. A VAN HIELE LEARNING PROGRESSION FOR SECONDARY STUDENTS USING LVAR IN MATHEMATICS

#### a) Methodology of the learning progression

The current teaching experiments (Cobb & Steffe, 1983) are evolving as students' van Hiele learning progression analyzes non-routine, real-world problems in addition to student assignments from the problem-solving process. It is constituted from (a) a learning trajectory in quadrilaterals (b) a learning trajectory in fractals. The trajectory in quadrilaterals follows the structure of the DHLP created in my PhD thesis. The difference is in the objects, which in this case have been selected from the real world [e. g objects in museums, mainly archaeological, in Greece].

The teaching experiment involved 81 students aged 13-14, equally separated into three classrooms. Every sub-class included the same number of boys and girls and the same number of high- or low-achievement students at the beginning of the year. The study investigated (a) ways to foster students learning by hypothesizing what the students might learn (e.g. develop real-world problem representations, reasoning and problem-solving, making decisions and receiving feedback about their ideas and strategies) working individually or collaboratively (b) ways in which students develop abstracting processes through building linking visual active representations and (c) ways to develop students' van Hiele level.

I was the teacher and the instructor of the activities. I developed the instructional activities based on an analysis of the results of my PhD thesis, with regard to students' evolution of understanding on instrumental decoding when they construct quadrilaterals. We worked as a whole class, trying to develop a form of practice compatible with social constructivism (e.g., Wood & Yackel, 1990). I was actively involved with the children, encouraging small group cooperation both in and outside of class, without intending to show the process to complete the activity. I started the activity with a question; after the answers were given, I continued with sequential questions to

clarify the explanations or to help students with the cognitive conflicts. Then, I asked the students to complete the task in the paper-pencil environment and collected their work to see the level of understanding from the correct answers. After the evaluation of the students' work, I continued with follow-up activities in the DGS environment to help the children reconstruct the solution methods. After the intervention with GSP activities, the paper-pencil work was repeated to see the difference in the students' learning and understanding of the concepts. Indicative of students' wrong representations will be presented and a short report made of their mistakes and misconceptions.

*"The situations that children find problematic take a variety of forms and can include resolving obstacles or contradictions that arise when they attempt to make sense of a situation in terms of their current concepts and procedures, accounting for a surprising outcome (particularly when two alternative procedures lead to the same result), verbalizing their mathematical thinking, explaining or justifying a solution, resolving conflicting points of view, developing a framework that accommodates alternative solution methods, and formulating an explanation to clarify another child's solution attempt" (Cobb & Steffe, 1991, p.395)*

The complete study includes:

- ...a detailed procedural analysis of the situations, the involved problems, in addition the problems' conceptual analysis, instrumental decoding and learning targets (e.g., different solving strategies, formulas or figure's decomposition). This includes the recognition and demonstration of transformations (e. g, recognition and drawing of symmetry lines or demonstration of reflections, translations and rotations) using multiple contexts (e.g., graphpapers, a computing environment). Furthermore, is described the recognition and utilization of properties that belong to a class of figures (or a subclass) and description of the characteristics of shapes and their relationships.
- ...an example of a theorem's LVAR modeling process (e.g an LVAR modeling for the Pythagorean theorem).

Students' uploading of assignments was facilitated through the free open-source Learning Management System Moodle (Modular Object-Oriented Dynamic Learning Environment) (Dougiamas, PhD thesis, described at <http://www.moodle.org.nz/>).

#### i. Presentation and analysis of problems

For the design of activities I always had in mind: "What would the individual have to know in order to be capable of doing this task without undertaking any learning, but given only some instructions?" (Battista, 2011, p. 515).

Case A: The problem was presented modeled in the dynamic environment. In the modeled dynamic representation, emphasis was given to the features associated with mathematics (e.g., the modeling of a kite can be done by constructing a rhomboid that emphasizes the verticality of the diagonals, etc.), rather than to other characteristics (e.g., the material, color, etc.). The students were able to experiment with the software tools on the digital image and to visualize the properties of the shapes that they were not able to perceive in the static environment.

Case B: The problem was not presented modeled in the dynamic environment, but the students were prompted to manage the image as if it was perceived in the natural environment. The students had to construct a simulation of the problem in a static, digital, or other physical means as a model of the natural environment. They also had to manage the (digital or not) image to gain intuition about the properties of the shape.

According to Johnson-Laird (1983) the human beings understand the world through the representations of the world they create in their minds. Johnson-Laird (1983) argues

*“to understand a physical system or a natural phenomenon one needs to have a mental model of this system that will allow [...] the person who will build it to explain it and to predict about it” (p. 430).*

In essence, the image conversion of the natural environment in the dynamic environment is a result of a complex process on the student's part. The student has first to transform the verbal or written formulation (“construct a parallelogram” for example) into a mental image, which is to say an internal representation recalling a prototype image (e.g., Hershkovitz, 1990) that s/he has shaped from a textbook or other authority, before transforming it into an external representation, namely an on-screen construction. The student needs to explore the shape of the natural environment (e.g., properties of shapes such as its symmetry lines, etc.) and then construct the scale model. The digital image plays a supporting role in understanding the properties of shape but also can bring to the surface students' cognitive obstacles and, consequently, lead to errors. These errors are mainly due to their vH level. As a result, students may not have the capacity to recognize the figure's properties, and, generally, to develop the solution with deductive reasoning. Especially for the fractal activities, the experimental teaching was carried out on 18 students at different school levels, including activities (on different software pages with linking representations) with increasing degree of difficulty depending on the age-related level of students. No student that participated had previously processed the software, or any other related software. As it was verified henceforth at many points of process the students were

led to conclusions and formulations of definitions that had not been made known during their course of mathematics.

#### ii. *Student's mathematical knowledge*

In secondary high school, the students are taught the kinds of quadrilaterals, which they are asked to memorize. Most of students are able to recall only the basic relations regarding perpendicularity and parallelism of the sides of quadrilaterals. Furthermore, students construct parallelograms in static means using their traditional tools (compass or ruler), which only fulfill the visual criteria. In Greece, dynamic geometry is rarely used in high schools to facilitate the teaching and learning of geometry. As it is concluded the teaching of reflective symmetry (or symmetry by axis) and symmetry by centre in a DGS environment is not correlated with the notion of symmetry and particularly the students do not examine the notion of symmetry in relation to quadrilaterals. Furthermore, the students' difficulties in constructing a figure are due to their ignorance of the different thought processes involved in dynamic rather than static means. The knowledge of a figure's symmetry is essential for students. I distinguished a few types of obstacles due to student lack of competence in instrumental decoding (:i.e. this is to say an instrumental obstacle). In the current study, I have devoted enough time for the students to understand the meanings (for example, the notion of symmetry by axis and symmetry by center) through the dynamic geometry software. The kinds of transformations on which the activities are focused are reflection--which corresponds to symmetry by axis in static means, rotation--which corresponds to symmetry by center, and translation. The dynamic geometry system helped students to instrumentally decode the properties of figures, as we will see in the description of the activities.

## IV. DESCRIPTION OF ACTIVITIES

### a) *Situation First (Visual-Holistic Reasoning -- Visual-informal componential reasoning): Recognizing quadrilaterals and symmetry in real world.*

The aim of the activity was the recognition of quadrilaterals and the investigation of the symmetry lines of quadrilaterals in a real-world context. Our actions included three phases: a tour in the museum, the teaching in the class (including training in my e-class: the operation of the e-class to facilitate posting and downloading of material), and finally, the realization of the activity for the students within a predetermined time. Briefly, the students had to construct a shape using the figures' properties, in terms of its sides and angles. The description of the activity consisted of the following parts:

1. Recognize the kinds of shapes that you observe in the decorative pattern of the image below (see Figure 5).

2. What kinds of symmetry do you recognize?
3. Construct the axes of symmetry and the center of symmetry in every image.
4. Construct the same image using your ruler and compass.
5. Construct the same motif using a dynamic geometry software.



Figure 5 : Traditional Greek embroidery

The problem is important for the development of students' ability to translate among different representational systems. An essential combination of visual skills, representational competency, and mathematical reasoning is required to solve this problem. A further aim of the activity was the creating of a diagram representing a real-world object in which the students can connect the abstraction, the art, and the timelessness of beauty.

*i. Phase One: Visit to the Museum of Greek Folk Art*

Guides [students of History and Archaeology at the University of Athens] took the students to different areas of the museum exhibits. The exhibits included local folk-costumes of various parts of Greece (e.g. mountainous Epirus, Thessaly, and the Aegean). These acquired meaning through the detailed presentation of the guides, who aimed to underline the particular characteristics of the local folk-pieces. The students were impressed by the embroidered women's costumes. Some had geometrical recurring motifs and expressed the inner desire of every woman [every bride] to have good fortune, happiness, and longevity.

Then, the students subsequently had to capture a part of the entire plan on paper.

*"On the ground floor of the museum, visitors will see elegant examples of traditional embroidery from the whole of Greece. They include polychrome*

*and white embroideries-laces and gold embroideries intended to meet the needs of dress, house and church. Particular interest attaches to pleated embroidered chemises of Crete, the relics of a female dress type with Renaissance roots that is found in other islands in the Archipelago during the period of Frankish rule (17th – 18th c.). Their hems are embroidered with alternating representations of gorgons, double-headed eagles, flower-vases, fantastic birds, and etc.*

(Excerpt from the text written in the description of the Folk Art museum website available at <http://www.melt.gr>).

I asked them questions such as: "What shapes can you 'see'?" "What kind of symmetry do you recognize in the decorative pattern?" The dominant feature of the costumes' geometric motifs [converted into images for the students' work] was the symmetry of its parts. As we know, the relations between depicted objects in a picture or additional information concerning the objects (e.g. colors or other symbols that convey a certain message) and their style allow us to place it in context. However, in a picture, the data could hinder students' ability to 'see'/(meaning perceive) the geometrical shapes/figures. For example in Figure 5, the symmetries in the pattern are apparent (central symmetry or axial symmetry). Additionally, it is also clear [to teachers] the symmetries of the shapes that form the overall motif. However, this is not true for students.

From the work of students resulted in the following conclusions: Regarding the functionality of the e-class, there was no particular difficulty with the operation of asynchronous learning by students. As to the concepts found that: students were not aware of the concepts of central and axial symmetry, did not understand the differences between quadrilaterals and for this reason they didn't 'see' the usefulness of such an activity, as some even use rice paper to replicate the project. In other words, it was found that students were not 'seeing' mathematics to the real environment and faced more difficulty in manufacturing patterns [see Figure 6]. Thus, I utilized this alternative way of teaching when I understood that students faced problems in understanding the concepts.

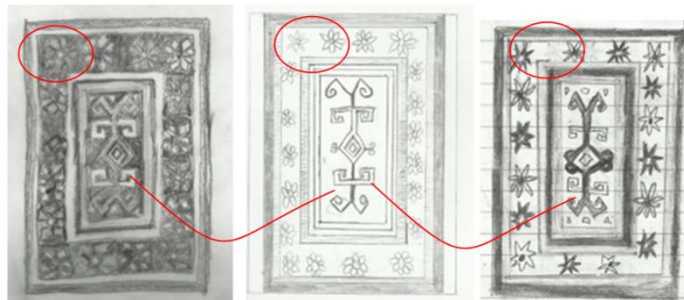


Figure 6 : Indicative of students' [pre-]constructions in the paper-pencil environment.

ii. Phase Two: Classroom [dynamic] organization

Firstly, the students recognized the parallel lines and parallelograms in the Figure 5. The students constructed the parallelograms using the “copy-paste” tools of the software or joined four line segments so they produced rectangular figures. Students make mechanical use of the software, which makes it impossible for them to understand the logic underlying the command options. It was my intention to familiarize the students with the software, “‘step by step’, in parallel

with the corresponding theory” (Mariotti, 2000, p. 41).

In order to construct a parallel line using the software, one has to select two objects: a straight object (for example a line) and the point from which the line parallel to the initial line will be drawn. Most students at van Hiele level 1 were unable to understand the sequential apprehension of the tools selection, because they were unable to understand the logic of the sequence of actions or unable to link this logic with the theory of geometry.

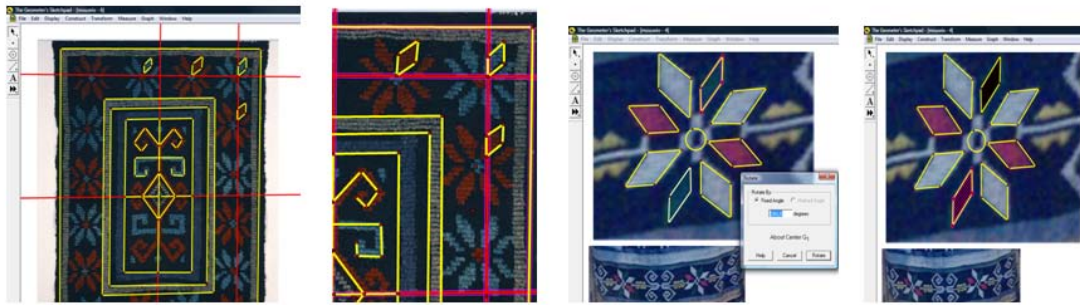


Figure 7 : Snapshots of the copy-paste input process in the DGS environment

For example a student (van Hiele level 1 at the pre-test) faced an instrumental obstacle which depended on the sequential apprehension of the objects to be used for the construction. The student tried to construct a parallel line by selecting the line alone and then the menu command, which is to say the student followed an irrational sequence of actions. At this point, s/he faced an instrumental obstacle and commended in an informal way on the non-activation of the software's command. Subsequently, student's interaction with the software, led to a cognitive conflict which helped him/her to apprehend the sequence of actions. Therefore, is the construction that leads students to “shape” inadequate or alternative definitions regarding parallelograms. The definitions followed the introduction of the parallelism and dragging tools of the software.

Then, the students participated in the process of introducing the concepts of symmetry through dynamic geometry software. In the lesson that followed, transformations were introduced to students using the GSP software tool. For example, we focused on the transformations we have to apply to a triangle to construct a parallelogram. In the images below, a mode-A LVAR construction in the software presented the translation of a triangle, the use of a coordinate plane, the rotation of the triangles (and the rotation angle that remained stable at every point in the triangles), the reflection of the triangles (including the visualization of the similarities and differences on the coordinates of the transformed images).

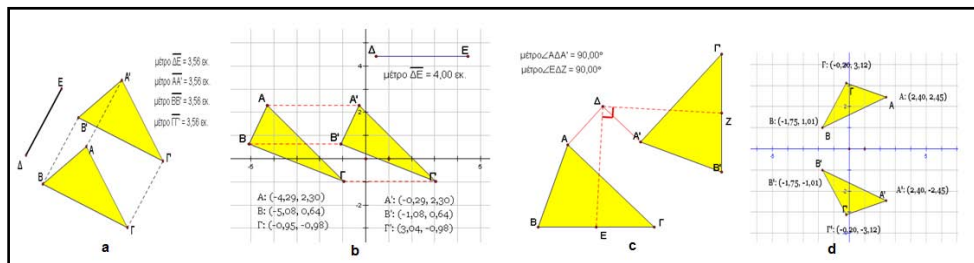


Figure 8 : Transformations and use of coordinate planes in the DGS environment (Patsiomitou, 2010b, in Greek)

The process of implementing the GSP software was simple, and students participated with great interest, answering my questions. Furthermore, the students demonstrated the capability to recognize the constructed parallelograms (e.g AA'B'B, AA'Γ'Γ in the figures 8 a, b). They globally recognized the axes of

symmetry of the image and the (sub) axes of symmetry in the subfigures (Figure 8).

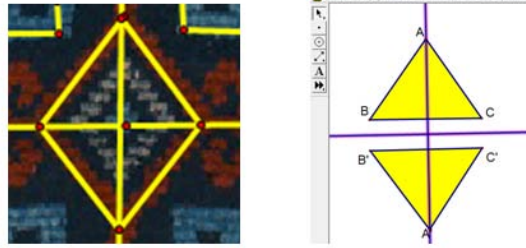


Figure 9 : Simulation of the object in the DGS environment

Then the students' task was to construct a rhombus based on the figure's symmetry (Figure 9). They dynamically reinvented that a single diagonal can divide the rhombus into two congruent isosceles triangles. Therefore, two congruent isosceles triangles can be together to form the shape of a rhombus. In Figure 10, the importance of the software's tools [e.g., the reflection tool] for the students' modification, or change their way of thinking, is represented. The most important indicator was that students tried to construct the symmetry by center of an arbitrary point on screen by using the reflection tool.

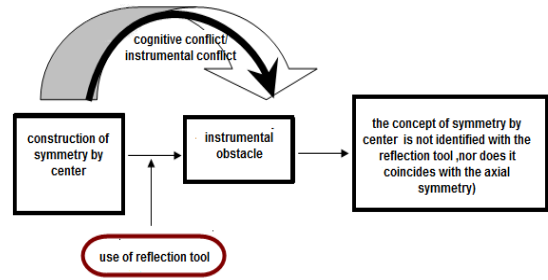


Figure 10 : A Pseudo-Toulmin model for the representation of the instrumental conflict

The utilization of the reflection tool during the previous phase led students, through instrumental genesis, to construct a utilization scheme for the tool. In this case, the students used the reflection tool by economy (Rabardel, 1995), despite having the option to use the rotation tool, in order to avoid the efforts required to use a less familiar one (Docq & Daele, 2001, p.200).

This action led to an instrumental obstacle as the result of the students' cognitive conflict with regard to the meanings of symmetry by center / axial symmetry.

iii. Phase Three: Students' meta[constructions]

The image below illustrates three indicative representations of students after the interaction with the DGS representations. The students dynamically reinvented meanings and the difference between terms [i.e. the meaning "axis of symmetry as perpendicular bisector of the segment which joins the original point and its reflected point", the difference between axial symmetry and symmetry center].

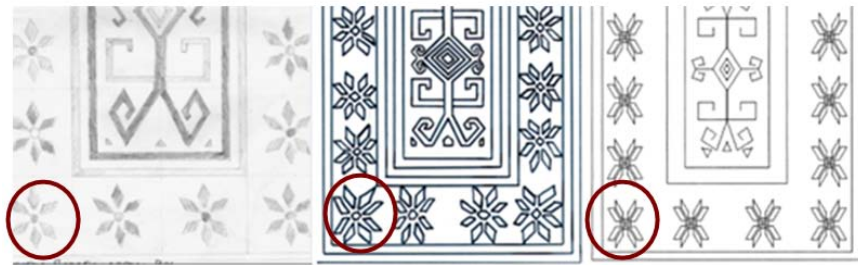


Figure 11 : Indicative of students' [meta-] representations in the paper-pencil environment

Students visualized the figure's symmetry, identified the symmetry axes of the quadrilaterals, but their representations [still] indicate that they recognized the shapes holistically. For instance, they globally recognized the sub figures of the flowers in the pattern, which were different for every student.

b) Situation Second (Visual-informal componential reasoning): Quadrilaterals' transformations

i. Phase One: Visit to the museum of Ancient Agora-Stoa of Attalos

Our actions included among others the visit to the museum of Ancient Agora. "The museum of the

Athenian Agora is housed in the reconstructed Stoa of Attalos, [...] Ever since, all finds from the area, together with the excavation archives, have been housed in the building, making the Stoa a significant research center for scholars and students of archaeology from all over the world"

(Excerpt from the text written in the description of the Athenian Agora website available at [http://www.ascsa.edu.gr/pdf/uploads/CATALOGUE\\_6\\_2\\_012\\_C\\_Layout\\_1\\_FINAL.pdf](http://www.ascsa.edu.gr/pdf/uploads/CATALOGUE_6_2_012_C_Layout_1_FINAL.pdf))



The archaeologists also discovered amazing mosaics in an ancient room. The next activity included students' construction of the mosaic, using their rulers and compasses, and the construction of the same motif using dynamic geometry software. In the representation of the mosaic the students tried to discover the angle of rotation of the parallelograms but they [still] confused the shape of parallelograms with the shape of rhombus because of their orientation. They tried to construct the successive parallelograms, but they failed to find the angle of rotation.

ii. Phase Second: The utilization of trace tool - reflection tool

The trace tool in correlation with the reflection tool proved essential for students' understanding of concepts; it helped the students to develop argumentation with regard to the equal distances of the

points (original-, reflected point-image), and the identification of the axis of symmetry as a perpendicular bisector. The construction of the figure was completed using the reflection tool of the software. Furthermore, the students discovered the axes of symmetry of the shape. They also considered what the center of symmetry is and in what angle the parallelogram can be reproduced. Moreover, students identified that a perpendicular constructed from the symmetry centre crosses the figure. Then they (1) recognized that the interior figure was a rhombus (2) if a line is perpendicular to one of the two parallel lines, it is perpendicular to the other (3) a perpendicular constructed from the symmetry centre of a rectangle to a side of the rectangle, crosses the midpoints of this side and its opposite side (4) the lines through the midpoints of two opposite sides of a rectangle dissect the rectangle into four rectangles that are congruent to each other.

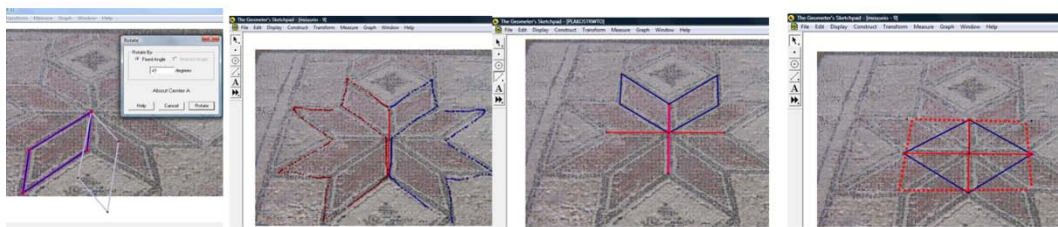


Figure 12 : Snapshots of the process in the DGS environment of a mosaic in the Ancient Agora of Attalos

The pseudo –Toulmin model below represents the process and the tools that led the students to construct the meanings.

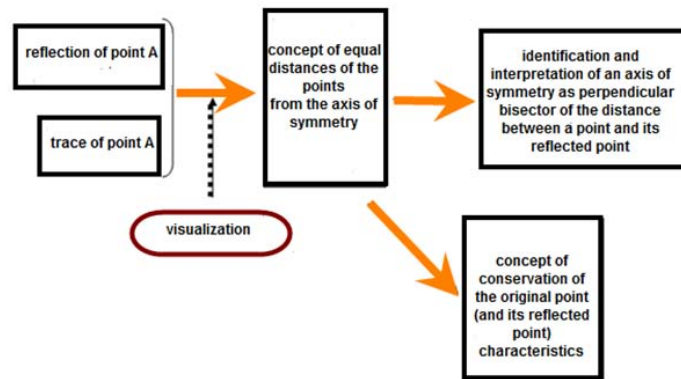


Figure 13 : A Pseudo-Toulmin model for the representation of the construction of the meaning

Through visualization of the object, students are initially led to use empirical methods (such as trace of points) in order to confirm that symmetrical objects lie at an equal distance from the axis of symmetry. Consequently, the students discovered a few properties of the figure, meaning the symbol character of the figure. Therefore, we had a theoretical construct which derived through instrumental decoding of the dynamic diagram.

c) Situation Third (Informal and insufficient-formal componential reasoning): Quadrilateral's properties in real context.

i. Phase One: Visit to our school courtyard

The modeling of the courtyard area of the school through static or dynamic means on the part of the students is an example of a real-world context modeling problem. The organization of the activity

consisted of three phases: (1) the processing of the courtyard pattern of the school during instruction, (2) measuring the tiles forming the courtyard's shape, and (3) modeling by students during fixed intervals. This problem is important for developing the capacity of students to convert between different representational systems.



Figure 14 : The courtyard area of the school

to measure the dimensions of the tiles. The aim was to construct a representation of the pattern to scale.

Skilful combination of visuospatial ability and representational capacity is required, as well as the capacity for mathematical thinking. The questions that I posed to students (e.g., "Is the inner quadrilateral a square or a rhombus?") focused on the recognition of the kind of quadrilaterals representing the exterior and the interior shapes on the floor and to calculate the area of the surface covered by the red and the white tiles.

Students' responses led to extensive dialogue/debate among them and gave me feedback. This example illustrates Cobb, Yackel & Wood (1992) claim that

*"students will inevitably construct the correct internal representation from the materials presented implies that their learning is triggered by the mathematical relationships they are to construct before they have constructed them. (Cobb, 1987; Gravemeijer, 1991; von Glasersfeld, 1978). How then, if students can only make sense of their worlds in terms of their internal representations, is it possible for them to recognize mathematical relationships that are developmentally more advanced than their internal representations? (p. 5).*

The conceptual frame mentioned was the areas of surfaces, the surface measurement units, conversion between different surface measurement units, and the areas of shapes. Students were required to consider the type of triangle formed by the diagonal of the square and then to justify the measurement of the formed angle. The measurement of the surface could occur in many ways (e.g., measuring the tiles forming the shape) in which the students had to observe the shape that each of them had and determine the area. The students used their geometric instruments (e.g., ruler, compass)

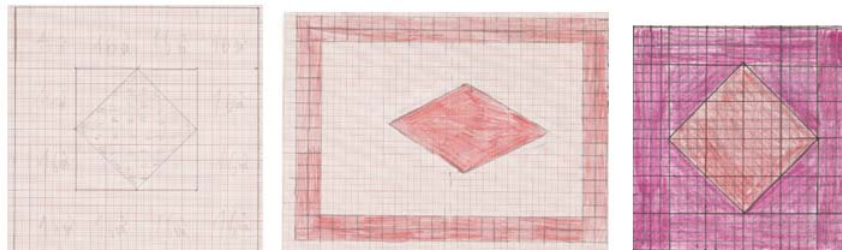


Figure 15 : Students' indicative constructions before the interaction with the GSP activities

The images (Figure 15) are indicative of students' representations in the paper –pencil environment. The students faced many difficulties connected with their conceptual obstacles in regards to the meanings of quadrilaterals. For example, many students perceived a rhombus rather than a square in the internal figure, due to the figure's orientation. Other students had constructed, in the middle, a rhombus whose edges coincided with the midpoints of the external figure.

#### ii. Phase Two: [Students'] 'dynamic' actions

The figure 16 represents a draft of the instructional design of the activity, using the software's tools. In the lesson that followed the in-class simulation, transformations of a lattice/grid were introduced to students using the GSP tool. While investigating the problem, the students used the rotation tool to rotate a segment by 90 degrees. Through this, they constructed the meaning of the right isosceles triangle. Moreover, the lines through the midpoints of two opposite sides of a square dissect the square into four squares that are

congruent to each other. These properties are elements of the object being built with the GSP tool.

The manipulation of the dynamic objects in the software led the students to construct the properties of the square, while the transformations of the dynamic objects led to acquire the symbol character.



iii. Phase Three: Students' representations of the courtyard

The students used graph paper and static or digital material of their choice (e.g., cardboard or dynamic geometry software) to represent the construction (Figure 19). The application was of particular interest, and students were able to calculate

the figure's area in several ways: (1) using as a unit the area of a tile and then that of a square (whose side was made of four sides of a tile), (2) calculating the area of the shape of a tile and from this, the area of the whole shape, and (3) calculating the dimensions of the shapes (squares) and then using geometrical formulas for the total area of the figure.



Figure 19 : Students' indicative constructions after the interaction with the GSP activities

This means that the students have developed thinking processes and applied skills, developing a mathematical model to interpret the realistic problem. They connected visual and formal criteria which is crucial for the transition from the lower levels, to the upper levels. Therefore, the information regarding the figure's properties were transformed to a signal.

characteristic feature of the museum is that the exhibits are all closely connected with the Athenian Democracy, as the Agora was the focus of the city's public life". (Excerpt from the text written in the description of the Agora museum website available at [http://odysseus.culture.gr/h/1/eh152.jsp?obj\\_id=3290](http://odysseus.culture.gr/h/1/eh152.jsp?obj_id=3290)).

d) Situation Fourth (Sufficient formal property-based reasoning): Quadrilaterals' structural analysis.

i. Phase One: Museum of the Ancient Agora -Stoa of Attalos.

"The Agora museum is housed in the Stoa of Attalos, a reconstructed building of around 150 B.C. The



Figure 20 : Snapshot of the Ancient Greek Geometric Pottery in the GSP dynamic geometry system.

The students' activity included the representation of the figures of Greek geometric pottery. First, the students recognized the shapes, and then they used their geometric instruments to construct the figure, but they faced difficulties to analyze the figure into subfigures.

parallelogram's symbol character was completed with its primary properties. In the activity aforementioned, the students recognized the parallelogram on the screen from the structure of its bisected diagonals.

ii. Phase Two: The visualization of parallelogram's diagonals in real world images

Simulations of a scissor lift or Centre Pompidou's designs (Figure 21) in the GSP have been introduced to students, in order to focus on and interrelate the meaning of a parallelogram with the bisection of its diagonals. This means that the

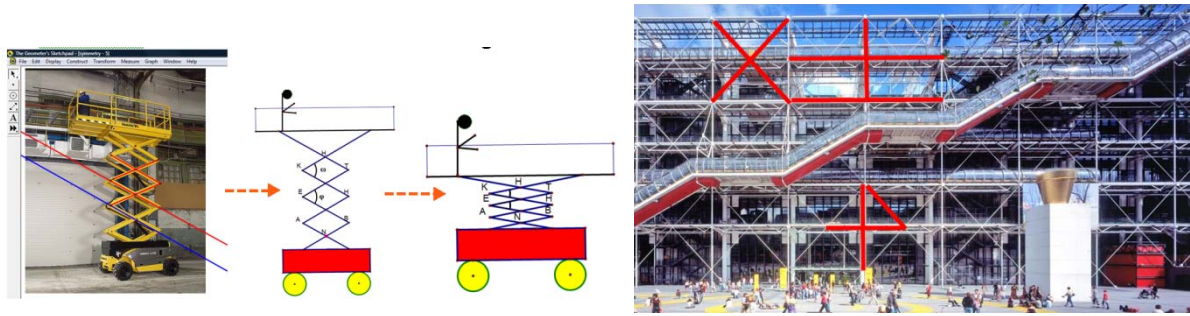


Figure 21 : The diagonals' visualization in real world images

iii. Phase Three: The instrumental decoding of the reverse process.

The instrumental decoding of the reverse process (i.e., the construction of a square) was more difficult for the students. The next step was the analysis of the relationships between parts of pottery's figures in the DGS environment. For example, the quadrilateral constructed from the connection of the points that intersect the diameters on the circle is a square (Figure 22). The students were not able to justify why the shapes were squares. They also changed the orientation of the diagonals in the DGS environment, applying the experimental dragging tool.

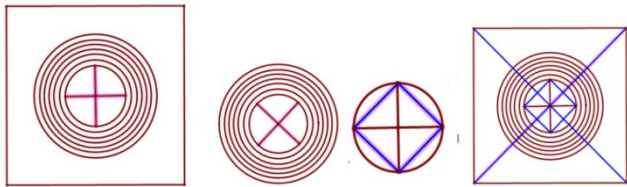


Figure 22 : Snapshots of the diagonals' transformations in the DGS environment

My questions helped them to identify the figure's properties and to analyze the figure into subfigures. They had to reverse the process, meaning they had to replace the figure with its properties. In other words, they had to construct the square's signal character. The sequence of questions led students to think of figure similarity. (For example, "Are squares similar figures?", "Are rectangles similar figures? Explain your answer"). Moreover, the students had to connect the meaning of the symmetry by center with the meaning of the segment's midpoint.

The students constructed the figure by taking into account the structure of its diagonals. They constructed two perpendicular lines intersecting at O, constructed a circle with center O and connected the four points where the circle cut the lines. It is crucial for the students to recall the properties of the figure's diagonals that were investigated in the previous phases of the research process by mentally linking the reverse representations in this procedure.

The students also used the custom tool 'symmetry' to reverse the process. The utilization of the

custom tool 'symmetry' twice with the second application point at the symmetry center O, will lead to the construction of two segments that have the same midpoint. Consequently, the meaning of "diagonals are dichotomized" can be constructed by the students through the use of the custom tool. Dragging the construction from a point-vertex, the properties remain stable, meaning point O remain the midpoint of both the segments. The students are able to recognize that: "if the diagonals of a quadrilateral have the same midpoint then the quadrilateral is a parallelogram or if the diagonals of a quadrilateral bisect each other then quadrilateral is a parallelogram". Subsequently, the students are able -by using the custom tool "symmetry" to transform an iconic representation into a verbal one through mental transformations.

"This is a very complex process since the students must have both conceptual and procedural competence, meaning the competence to instrumentally decode their mental representations of a set of properties with actions through the use of tools. This means, for example, to interpret the congruency with the circle tool and simultaneously bisect with the custom tool. Furthermore, for them to construct the hierarchical categorization and definition of figures through their symmetrical properties and in accordance to their understanding." (Patsiomitou, 2012b, p. 71).

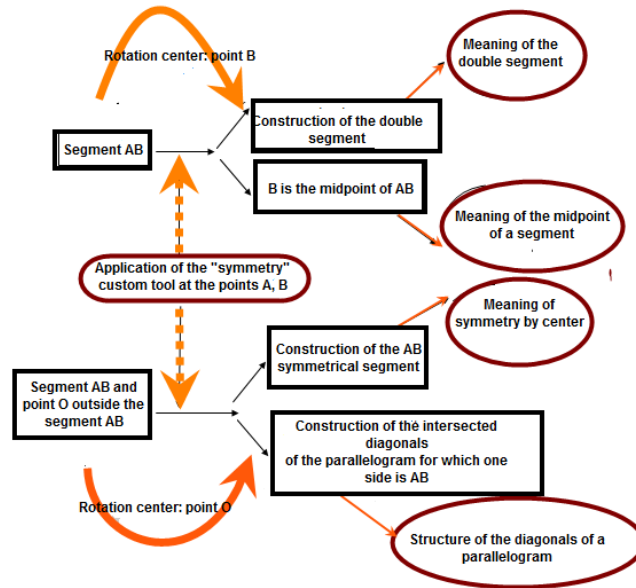


Figure 23 : The utilization of the custom tool 'symmetry'

The process is described in the pseudo-Toulmin's model above (Figure 24). The diagram expresses the way in which students in cooperation constructed the square, using the Sketchpad tools. Through the construction, they extended the structure of

the intersected diagonals, including the meaning of the perpendicularity and the congruency: "[a square's diagonals] are perpendicular and congruent segments intersected in a [common] midpoint."

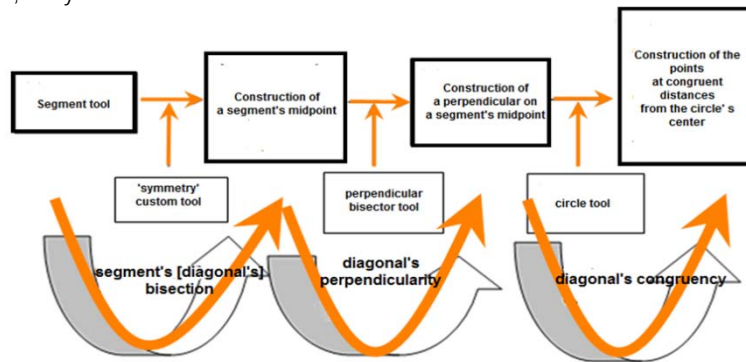


Figure 24 : The instrumental decoding of the reverse process

In the Figure 25 (Patsiomitou, 2012b, p.72) we are able to observe the linking representations of the diagonals of different types of parallelograms. Dragging theoretically the endpoint of the diagonals of the parallelogram in order these to acquire the property of the perpendicularity leads to the structure of the rhombus diagonals (or a square's diagonals). Dragging theoretically the endpoint of the diagonals of the parallelogram in order these to acquire the property of the congruency leads to the structure of the rectangle's diagonals. The construction of two arbitrary diameters in a circle (i.e. the diagonals are not perpendicular to one another) leads to the structure of the diagonals of a rectangle. The construction of two diameters perpendicular to one another in a circle leads to the structure of the square's diagonals. In this way conceptually and procedurally linking representations

are created.[...]. Subsequently, this learning path can lead to the development of an abstract way of thought through the development of linking representations in student's mind. .

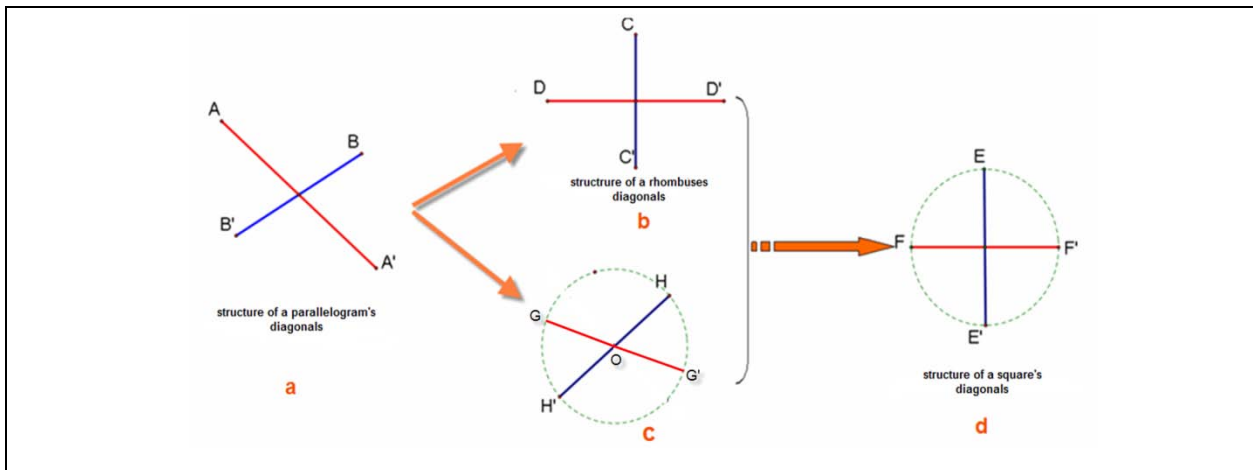


Figure 25 : Linking Visual Active Representations of parallelogram's diagonals (Patsiomitou, 2012b, p. 72)

e) Situation Fifth (Relational –Inferential Property- Based Reasoning- Empirical relations). A structural process  
 i. Phase One: Building structures

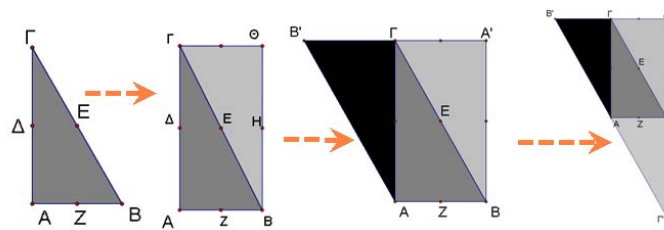


Figure 26 : A structuring process (Patsiomitou, 2010b, in Greek)

The next activity was the transformation of a geometrical figure that was used as a building unit for the construction and a means by which the students could construct the meanings of theorems inductively and experimentally, which were included in their class

curriculum. For example, in the figure 26 the rotation of a right triangle leads to the construction of a rectangle, then to the construction of a trapezoid, and, finally, to the construction of a right triangle whose sides are double from the sides of the initial triangle.

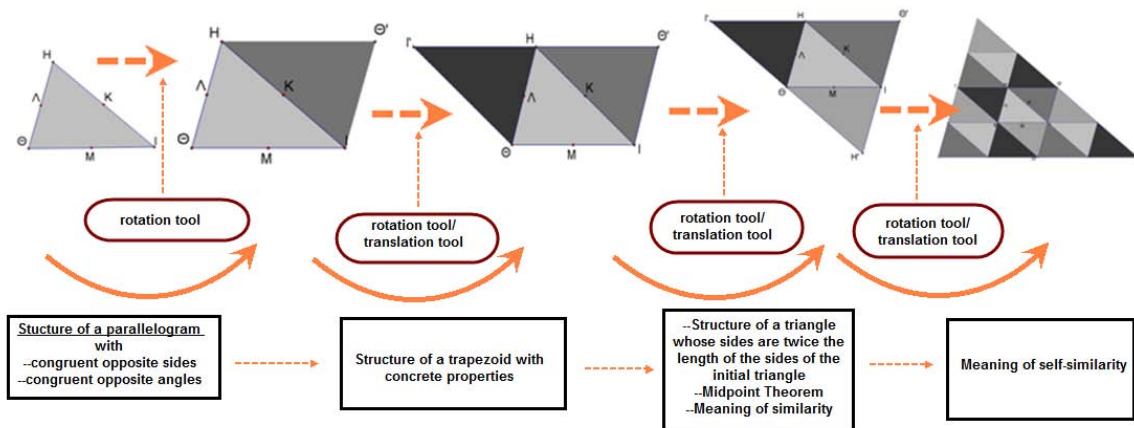


Figure 27 : The structuring process of the meaning of self-similarity (Patsiomitou, 2010b, in Greek)

In the figure 27 the rotation of a triangle leads to the construction of a parallelogram, then to the

construction of a trapezoid, then to the construction of a triangle whose sides are double from the sides of the

initial triangle, etc. Most important is the development of students' correlation of the properties (for example, "How is the meaning of a right and isosceles triangle linked with the meaning of a rectangle, and how is this consequently linked to the meaning of a square?", "What are the similarities and differences of the properties of a square and a rhombus, etc., as a result of the different structuring of its figure?", "How do the similarity of the building block's figures affect the similarity of the sequential figures?").

The students organized their thoughts for the sequential steps of the construction (for example, "What should be the property that must have a right triangle to be the building unit for the construction of an equilateral triangle?" or "What are the properties the sequential figures have?") The study of the building block's properties helped the students to organize the properties of the figure evoked from the initial figure. This process is in accordance with what Freudenthal (1973, 1983) has told that the teaching and didactic process must focus at the understanding of the structuring process and not the learning of ready-made structures. Moreover he argued that students could discover mathematics when they work with contexts and confront interactive and reflective activities.

ii. Phase Two: The Pythagorean Theorem through LVAR representations

In their calculations, the students had to use the Pythagorean Theorem. For this, the next activity was aimed to increase understanding of the application of the theorem in the class.

The teaching process consisted of three items:

- First, the visual proof of the Pythagorean Theorem with the utilization of linking visual active representations that I created using the Geometer's Sketchpad.
- Second, the meaning of the Pythagorean Theorem, and generalizations of the concept.
- Third, the extension of the Pythagorean Theorem to fractal structures (e.g., the construction of Pythagorean trees), such as successive calculations, the areas of squares, etc.

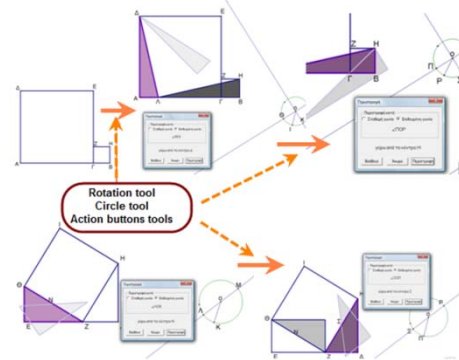


Figure 28 : Visual Pythagorean Theorem's proofs (Patsiomitou, 2010b, in Greek)

Students participated in the process of introducing and exploring the Pythagorean Theorem. The instrumental orchestration process (e.g., Trouche, 2004) included a laptop computer and an interactive board. Initially, the students were guided to explore the Pythagorean Theorem visually. Then, I asked the students to construct the shape using their paper and pencil environment as an assignment in class to determine how they perceived the objects.

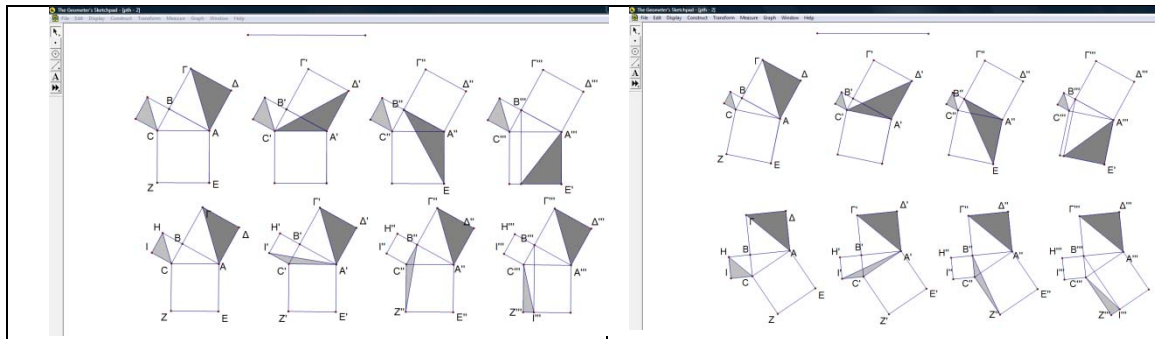


Figure 29 : LVAR translations of Euclid's proof (Patsiomitou, 2010b, in Greek)

Then the students conducted a visual proof of the Pythagorean Theorem with LVAR representations (e.g., Patsiomitou, 2010a, b) that I was created using the Geometer Sketchpad software. It captured the interest of the students. The combination of the interaction techniques of the software (e.g., navigation links, buttons to hide/show objects) for the production of

visual mathematical representations (VMR) (Visual Mathematical Representations) (Sedig & Sumner, 2006) can lead students to develop conjectures, analyze the problem, and synthesize the solution.



In the images of figure 29, we can see the linking representations of the Euclidean proof of the Pythagorean Theorem. The successive phases of the constructional steps have been achieved using transformational processes like the use of the translation command (Figures 29). By dragging a point of the original configuration or the translated images, the students can observe the processes that emerged previously being modified simultaneously. Students are able to directly assume or infer the properties and the interrelationships between figures from properties indicated on the diagram by conventional marks (for example the equality of angles, or the angles measurements). In the first row, four linking [translated] representations led the students to understand that the half square is transformed to the half rectangle. The same is visually demonstrated in the second row for the other square. The important point from the LVAR constructions is that the students can transform the shapes simultaneously and see the same theorem from a different orientation. Additionally, an important point,

segment, or shape is highlighted as the students develop their explanation orally.

I explained to the students that this method provides a visual confirmation of the Pythagorean Theorem and pointed out the need for proofs. The challenge is the interaction of students with LVARs to help them develop their level of geometric thinking. A pupil can develop his/her level of knowledge by proceeding through increasingly complex, sophisticated and integrated figures and visualizations to a more complex linked representation of problem, and thereby moving instantaneously between two successive Linking Visual Active Representations only by means of mental consideration, without returning to previous representations to reorganize his/her thoughts (e.g., Patsiomitou, 2008a, 2010a; Patsiomitou & Koleza, 2008). A student voluntarily presented the other students with the dynamic objects and the transformation of the shapes, which was a part of the process. If someone failed to provide the correct answer, the other students tried to help, expressing their point of view.



Figure 30 : Student's construction of the Euclid's proof linking representations (see [2, 4])

The action involved the modeling of the Pythagorean Theorem using a cardboard. The students' understanding emerged because of their interaction with

the LVAR dynamic representations within the community of practice in the mathematics classroom.

iii. Phase Three: Transformations and calculations of geometrical objects

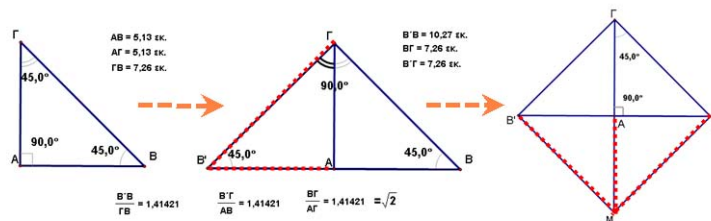


Figure 31 : Transformations and calculations of a right and isosceles triangle

LVAR played a significant role in developing students' deductive reasoning is clear from the fact that students demonstrated a shift from visual to formal proof. For example, in the figure 31 the rotation of a right and isosceles triangle leads to the construction of another right and isosceles triangle with concrete properties. Recognition of structures leads students to schematize and manipulate mental objects. Therefore,

construction through sequential transformations can lead to the organization of the figure's properties. This process is more than the synthesis and extension of properties.

iv. Phase Four: Selfsimilarity and calculations

The image (Figure 32) represents a work carved in stone from the famous workshop in Tinos Island. "Works carved in stone by popular craftsmen are used

either as independent ornamental structures (fountains, grave monuments) or as architectural features of practical and decorative purpose". (Excerpt from the text written in the description of the Folk Art museum website available at <http://www.melt.gr>)

Concretely, the image mentioned above is a snapshot of the image's input in the Geometer's Sketchpad dynamic geometry system. The repetition of the square's construction inside the figure leads to a fractal construction.

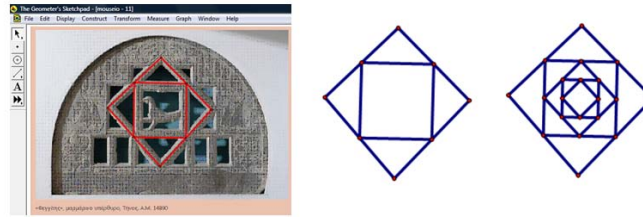


Figure 32 : Snapshot of the image's input in the Geometer's Sketchpad dynamic geometry system

The question was the following: "What kind of quadrilateral is shaped by joining the midpoints of the external quadrilateral?" For any quadrilateral, we can prove that the internal quadrilateral constructed by the midpoints of the sides of the external quadrilateral is a parallelogram. The students learn to prove this through a procedure of the application of the midpoint-connector theorem. In the image above, the interior figure is a square, as is the exterior figure.

If the exterior quadrilateral becomes a rectangle, then the interior--constructed by joining the midpoints of the initial--will become a rhombus, the next interior constructed will become a rectangle, etc.

The students can visualize a secondary property of the rectangle (for example that the axes of symmetry of the rectangle can be interpreted as

diagonals of the rhombus, in other words can be interpreted differently and acquire a second role. Then the symbol of rectangle is transformed to the signal of rhombus. It is what many researchers have discussed (e.g., van Hiele, 1986; Patsiomitou & Emvalotis, 2010a, b). The Toulmin's model diagram below is a representation of the way students expressed their thoughts. They told that "If the figure is a rectangle, then its diagonals are congruent, so these segments --that join the midpoints of the opposite sides-- are parallel and half the length of the diagonals".

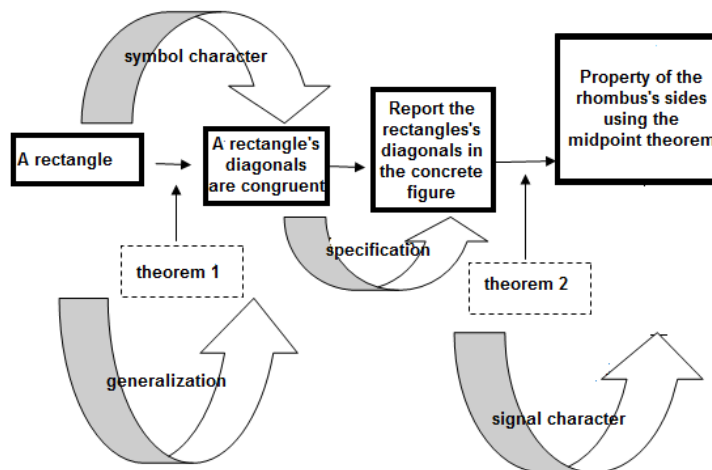


Figure 33 : A pseudo-Toulmin model for the representation of the way students expressed their thoughts

f) *Situation Sixth (Componential analysis). Modeling fractal objects using static and dynamic means*

i. *Phase One: Introduction to the world of fractals*

The 6th situation led students to think about self similarity, which is not included in high school

curriculum. The objective of the situation seventh was to awaken students in mathematics that are not included in their class curriculum. Moreover, because of the scheduled curriculum is difficult for them to explore. Martinez (2003) writes that "Mandelbrot coined the word "fractal" (from the Latin word "fractus", meaning

fractured, broken) to label objects, shapes or behaviors that have similar properties (self-similarity) at all levels of magnification or across all times, and which dimension, being greater than one but smaller than two, cannot be expressed as an integer" (reported in <http://www.fractovia.org/art/people/mandelbrot.html> ). The plan was to incorporate and illustrate fractal geometry --or facilitate the understanding of topics from geometry-- in already existing curriculum (e.g., fractions, proportion and ratio, calculations of area and volume, logarithms and exponentials, sequences and series, convergence of geometric series, geometry of plane transformations etc.) Furthermore, the enrichment with fractals into existing curriculum helps students to develop their imagination and apply mathematics outside the classroom, in real-world activities in cases that other students couldn't see the relevance. For example, among students' kites a highlighted one existed, constructed with Baravelle spirals (e.g., Chopin, 1994; Patsiomitou, 2005) fractals, of a student 12 years-old who participated in Fractal group.



Figure 34 : Students' constructions of kites

Mathematical concepts related to the construction and investigations of a fractal are divided into geometric and algebraic segments, which cover almost all concepts included in the high school curriculum. For the fractal constructions the Geometer's Sketchpad dynamic geometry software has been used which is the best dynamic geometry program for facilitating fractal constructions because of the in-depth iteration process that helps students gain strong intuition for the meanings (Patsiomitou, 2005, 2007).

The students watched videos exploring the fascinating world of fractals. The videos were posted, in the Moodle environment. The language of the videos was English, which did not cause dissatisfaction or difficulty for the students. Moreover, it is well-known that the language of mathematics is common internationally, and in the videos, common notations for mathematical concepts were presented. The students could also cooperate to collaboratively answer questions and complete a text for the golden rectangle, gathering information from websites or creating their own constructions. Mandelbrot or Julia fractals fascinated the students because of the beauty of the objects they observed. Some of the students processed natural fractals (e.g., broccoli, cauliflower) to understand that a fractal structure does not change. The shape and the size of the object do not affect the structure and the self-similarity of the objects.

### ii. Phase Two: Modeling fractal objects

The design of the activities and the experimental process that is reported here is an excerpt of my Master's thesis (Patsiomitou, 2005). This process has been repeated in the students' fractal group in the previous school year. For example: the construction of a "Pythagorean Theorem" custom tool, as well as the application of a "Pythagorean Theorem" custom tool recursively, led them to create Pythagorean fractal trees. Via the proposed activities we are able to investigate whether the construction of the fractals implementation or via the custom tools or of the process iteration can assist in investigating open-ended problems whose objective is the standardisation of intuitive ideas and the development of abstract processes. Moreover, we are able to investigate whether the students can be imported into the basic notions of infinitesimals and their use in calculus.

I conceived of LVAR representations when creating linked pages in Sketchpad files to construct fractals for my Masters thesis. Here is explained the rationale I followed in the design process. The most important parts of the design and research process are going to be mentioned here, enriched, to explain the importance of linking visual active representations, instrumental decoding, and RVR—as LVARs have been illustrated later (e.g., Patsiomitou, 2012a, b). The modeling and construction of an in-depth fractal structure is difficult or impossible with familiar geometry instruments (ruler and compass). Although the students' construction started in the paper-pencil environment, they felt it necessary to continue their construction in dynamic geometry software. The construction of the Sierpinski triangle fractal was one of the favorite subjects for the students. Moreover, the discussion expanded on the concept of a golden rectangle and golden spiral, and other spirals, such as the Fibonacci sequence, concepts that enriched the mathematical world of the students. Below I describe the way in which the students constructed a Sierpinski triangle in the DGS environment through two different ways: that of a custom tool (script) and thereafter application of the tool in (n) steps or the application of functional process of iteration (Steketee, 2002, Jackiw & Sinclair, 2004) to the initial construction (or even the composition of the two modes).

#### Phase Two -Part 1: The construction of the Sierpinski triangle via custom tools

For the construction of the Sierpinski triangle, the students started with an isosceles triangle (or an equilateral) and the midpoint of its sides (Patsiomitou, 2005, 2007). Then they guided to build a custom tool in order to continue the process. The students had to grasp the process in order to construct a Sierpinski triangle in-depth.

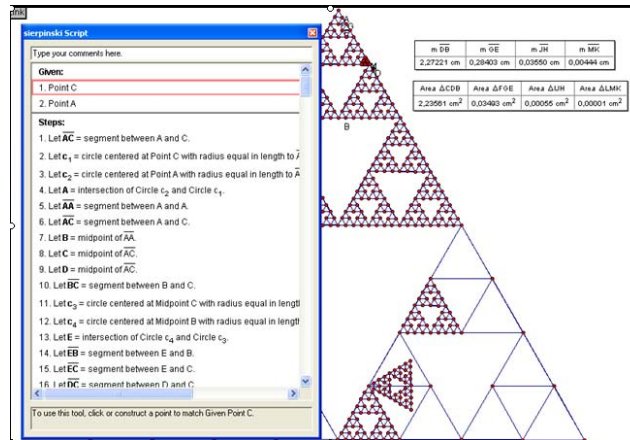
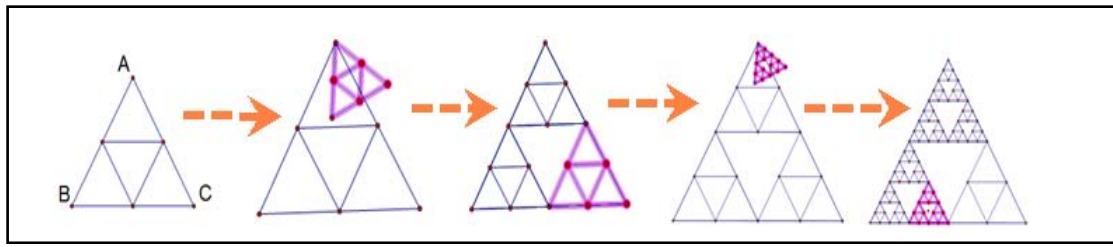


Figure 35 : Sierpinski's construction (Patsiomitou, 2005) via the utilization of sequential custom tools

The process is represented in the pseudo – Toulmin diagram below

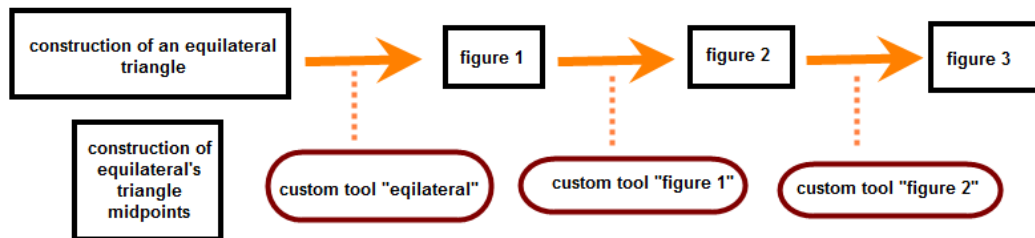


Figure 36 : A pseudo-Toulmin model for the representation of the process

From an instrumental genesis perspective, the students can construct an instrumented action scheme by using the custom tool, and then a higher order instrumented action scheme. Therefore, the custom tool 'equilateral' acts as a building unit in the genesis of the higher-order scheme, exactly as Drijvers & Trouche (2008) argue:

"The difference between elementary usage schemes and higher-order instrumented action schemes is not always obvious. Sometimes, it is merely a matter of the level of the user and the level of observation: what at first may seem an instrumented action scheme for a particular user, may later act as a building block in the genesis of a higher-order scheme. [...] a utilization scheme involves an interplay between acting and thinking, and that it integrates machine techniques and mental concepts [...] the conceptual

part of utilization schemes, includes both mathematical objects and insight into the 'mathematics of the machine' (p. 372).

The sequential creation of custom tools led the students to grasp meanings; however, most of the students had difficulties in understanding the structure of the triangle as the process evolved.

*Phase Two -Part 2: The iteration process*

In order to approach the task we constructed an equilateral triangle and from the midpoints of its sides the next equilateral and so on. The problem that we discussed concerned the calculation of the sum of the areas of the successive equilateral triangles in the interior of the shape. The whole iteration process can be demonstrated using Geometer's Sketchpad software to make it understandable for children aged 13-16. If we build a custom tool ("Area's sum," for example) that

finds the sum of the successive triangles, divide the sum with the area of the initial triangle, and repeat continually, we will get a result (e.g. 1, 25). The structural repetition of the triangles in-depth, as well as of the calculations, will not change the results. The next figures (Figure 37) demonstrate the linking of the visual active representations of the calculations, which generalizes the process. The final result is equal to 1, 3333..., meaning that the limit of the sequence of the infinite sum

of the areas approaches the 1, 3333...number as is strictly proven. The resulting sequence is formed by the sum of the areas of triangles (>8 triangles) made in each iteration. This means that we finally have a sequence of terms equal to 1.33333. In this way, the students understand that the size of the triangles does not affect the ratio of the sum of the area, which is approximately (~ 1.33) and remains stable, even if we continue the process.

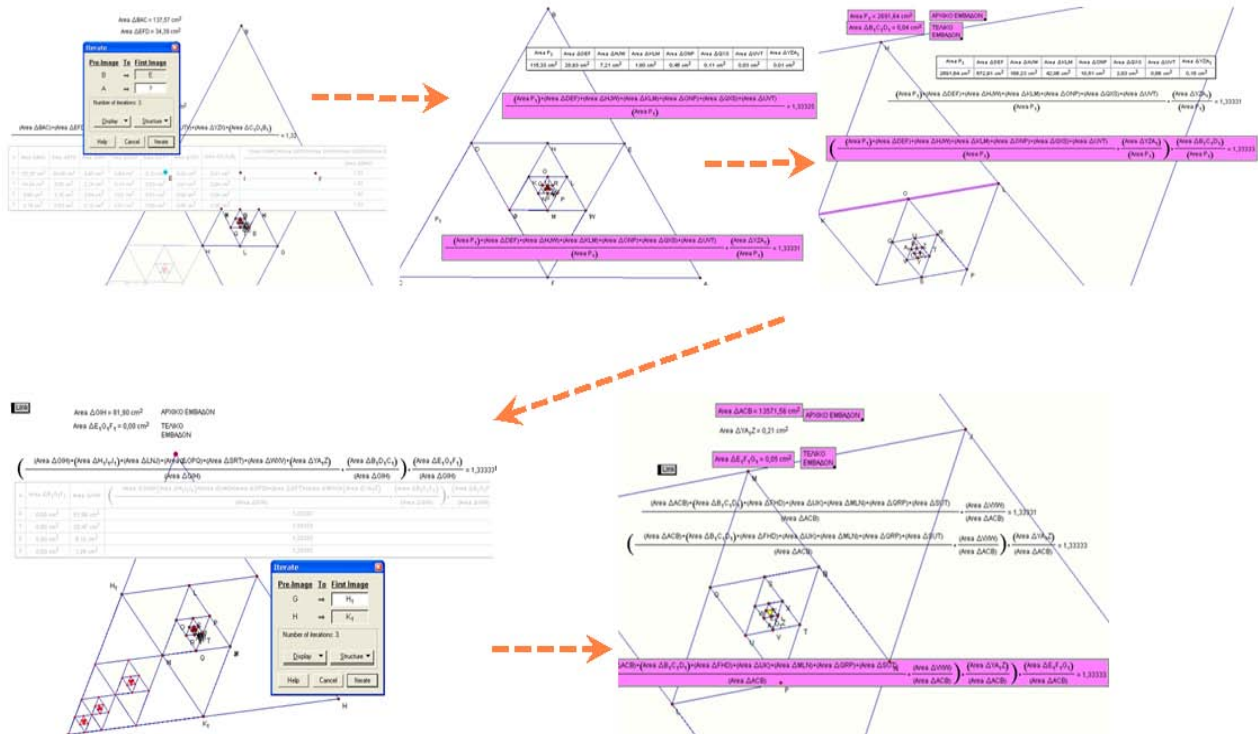


Figure 37 : Linking Visual Active Representations of Sierpinski's iteration process

How easy is it for a teacher or student working in the paper-pencil environment to create these representations with the software's accuracy or to synthesize all these together with precision and speed? From a mathematical perspective, we could mention the following:

If the area of the initial triangle is equal to E (the first term), every one of the triangles being built by joining its sides' midpoints has an area equal to  $\frac{E}{4}$ .

This series is geometric, with the constant ratio = 1/ 4. The question is about the calculation of the addition of triangles' areas in depth.

Meaning, the sum of  $E + E/4 + E/16 + \dots$  whose each successive term can be obtained by multiplying the previous term by  $r = 1 / 4$  ( $|r| < 1$ ). This infinite

sum can be calculated by applying the formula  $\frac{\alpha}{1-r}$  with a constant ratio between successive terms, or equally

$$\sum_{n=0}^{\infty} \frac{E}{4^n} = \frac{E}{1-r}$$

The sum is equal to  $4/ 3 = 1, 33333\dots$

$$\frac{1}{1-r} = \frac{1}{1-\frac{1}{4}} = \frac{4}{3}$$

iii. Phase Three: An experimental process with the Fractal's Group

The experimental process that is reported here is an excerpt of the paper "Fractals as a context of comprehension of the sequences and the limit in a Dynamic Geometry environment" (Patsiomitou, 2007)

included in the Electronic Proceedings of the 8th International Conference on Technology in Mathematics Teaching (ICTMT8) which took place in Hradec Králové (excerpt of my Master's thesis). The most important parts of the research process are going to be mentioned

here, enriched, to explain the importance of linking visual active representations, instrumental decoding, and RVR—as LVARs have been illustrated. This process has been repeated in the students' fractal group in the previous school year.

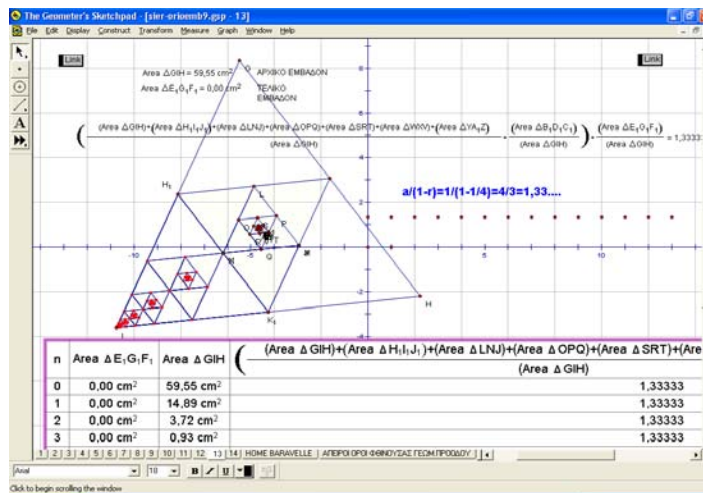


Figure 38 : The repetition of the number 1, 3333.

The team was constituted from 6 students 15-16 years old. The students at the school had not been taught about the sum of infinite terms of a geometric progression, because it was not included in their curriculum. The students initially observed that the areas in the interior of the shape were decreased by the ratio  $r = 1/4$ . This led them to the definition of the geometric progression for the areas of the shape and to the calculation of the sum of the infinite sequence  $E + E/4 + E/16 + \dots$  where  $E$  is the area of the initial triangle,  $E/4$  the area of the next internal triangle and so on. The inquiry process investigated if the students could perceive the meaning of the limit of the sequence of the infinite sum of the areas (approximate result almost equal to 1, 33333...).

The students through guided questions calculated the sums of the areas of the 2 first equilateral triangles and then divided it by the area of the initial triangle. Thereafter they calculated the sums of the areas of the three first triangles and divided this again by the area of the initial triangle. The process continued with the construction of the suitable custom tool that repeated this inductive process. When the process reached the 9 first steps the sum of the 9 internal repetitions of the areas of the equilateral triangles within the shape and the calculation of the ratio was stabilised at 1,33333 even when the initial triangle's shape was increased by dragging. Therefore, the generalisation of the process resulted from the process of iteration. With the assistance of the dilate tool and zooming into the depth of the construction thus dilating the structure the afforded impression was that of an infinitely continuous structure which had in actual fact remained unaltered

and constant. The students confirmed the repetition of the number 1,33333 on the table for (n) first steps of iterative constructional steps. In the beginning they applied the process and a shape resulted at the centre of the initial shape. Dilate tool assisted them to see into the centre of the shape and extend their mental representations.

[...] In the latter activity we were led towards the construction of a branch of the Pythagorean tree using the modes that were mentioned before. The students had not comprehended the graphic representation of the sequence when it had been discussed with static means in their class during their course of mathematics. Their initial reaction was to connect the isolated points that resulted from the plotting of the areas of the successive squares, in order to produce a continuous curve. This reaction of the students was a result of misconception of the definition of the domain of any sequence which is the natural numbers, but more so the result of the correlation of the graphic representation of the functions as it has been introduced by static means. The female student that with certainty repeated the definition of the sequence, but had comprehended this fine point about the graphic representation of any sequence, in no way displaying her misconception by inquiring about the fate of the intermediate points of the continuum.

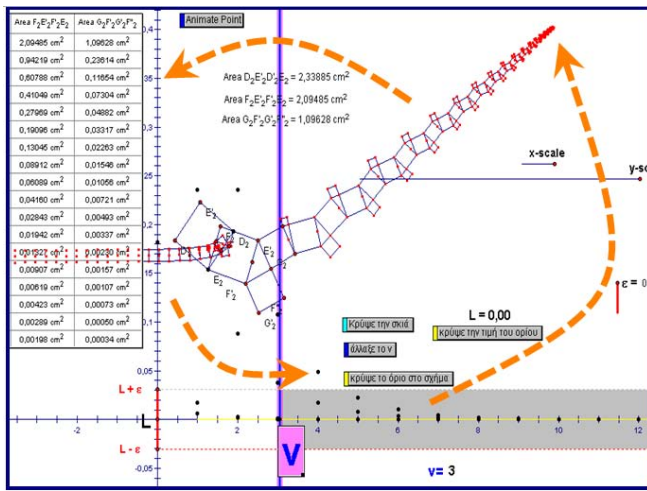


Figure 39 : Multiple linking visual active representations

The connection of the concept image with the concept definition of the meaning (Vinner, 1983) and finally the graphic representation was created through the environment of the software. At this point in the shape we have used as base for the development of the activity the file seqlimit.gsp. (Retrieved from <http://www.teacherlink.org/content/math/activities/sketchpadv4.html>). As it is described by the authors: “[The file has been designed] to help students graphically visualize the concepts behind the formal definition of the limit of a sequence. Given a value for epsilon, students can manipulate N to find a value for N beyond which all further terms of the sequence lie within the distance epsilon from the limit”.

In this sketch I had created an adaptation of the shape of fractal Pythagorean tree (Figure 39). The process of animation can produce the changes in the tabulated measurements (calculations) that allow the user to examine the dynamic process. These changes come as result of the fluctuations in the size of an artefact-fractal which have the possibility of increasing (decreasing) and altering orientation. The students consequently had an environment of multiple linking visual active representations in which the shape of the fractal had been linked with the table of the measurements via the functional process of iteration, which continuously could be linked with the graphic representation of the sequence. [...] (Patsiomitou, 2005, 2007).

As a result of the construction and application of the custom tool as much as the process of iteration the direct perception of the user is attained in regard to the steps in the development of the construction pertaining to (Patsiomitou, 2005, 2007):

- The repetitions in the measurements or calculations of the areas of initial shapes
- The developmental way of the construction of the shape and

- Its orientation towards the sequential steps of the construction on the screen's diagram or in successive pages of the same file.

Through the application of the custom tool the possibility is given to the user to acquire an inductive way of thinking for the finite steps of the construction but the generalisation with regard to the constructional result can be achieved from the process of iteration which inductively renders the construction theoretically to infinity (Patsiomitou, 2005, 2007, 2008d). This function of the software also constitutes a certain crucial and essential particularity, while the construction with a compass and a ruler as formal tools of static geometry has a beginning and an end. Moreover, “a teacher is able to distinguish different levels of acquisition and mathematical engagement with a fractal topic [as] scripts [ / iteration command] represent an abstraction of his/her own work or process, and thus using them as “abstract” tools require one level more advanced or sophisticated a conceptualization than using “literal” tools like the compass or straightedge.” (Personal e-mail communication with Nicholas Jackiw on September 29, 2005).

In the software, via the process of iteration we have the potential of the constructions thus becoming more complex being in theory rendered inductively to infinity. The result of the process of iteration is the construction of the tables that repeat the process of initial measurements and calculations in dynamic connection with the shape, thus increasing (or decreasing) the level of the process of iteration while the software adds (or removes) the next level of measurements (or even calculations), whereas in the first column of the table the sequence of the natural numbers is presented. In that way through this operation, the environment of the software promotes the exploration of the sequences and of the series. The iteration process by functioning thus has integrated or embodied the meaning of sequence while there is a direct connection between the user's perception and the abstract mathematical meaning (Patsiomitou, 2005, 2007). Therefore, I think that The Geometer's Sketchpad v4 [-v5] is the best tool to introduce fractals in classroom not only for aesthetic purpose rather than for the pursuit of their very interesting mathematics.

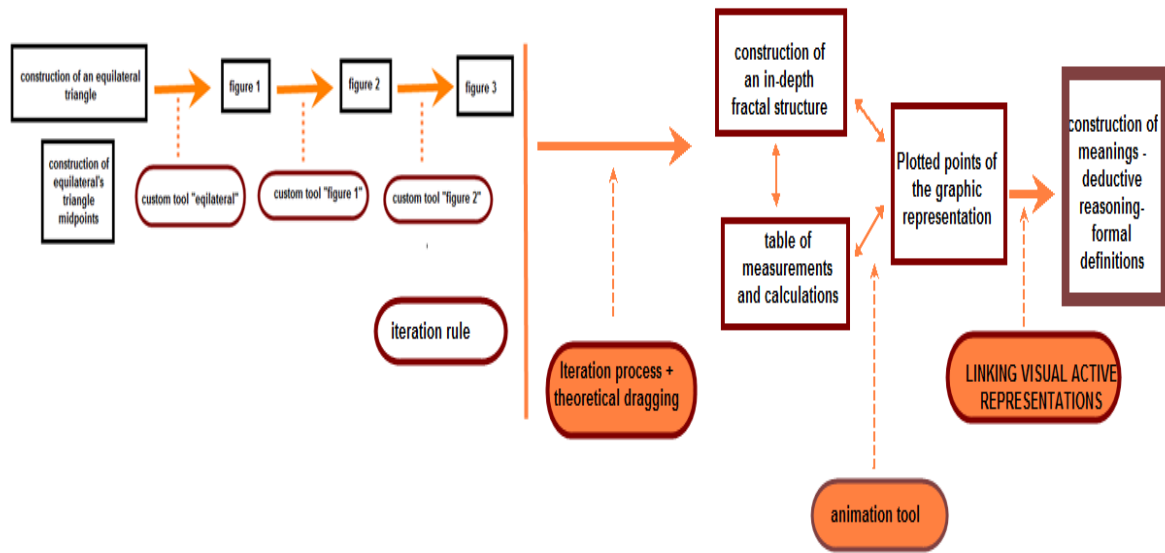


Figure 40: A Pseudo Toulmin model explaining the birth of the meaning 'Linking Visual Active Representations'.

The structures of fractals, [by applying the meaning of dynamic LVAR representations], aims that students (a) review most of theorems, (b) identify the potential weaknesses and cognitive obstacles that students face in their effort to understand the process, (c) develop the links between the virtual representations and the formulations with which students justify their construction, as a result of understanding the figures' transformations and symmetry, and (d) develop most of the competencies described in the beginning of the article and higher-order level skills (e.g., generalize patterns using recursion, use algebraic formulae and symbolic expressions to explain mathematical relationships, etc.) than those that they are able to develop through traditional mathematics. This is very important for their movement through vH levels.

## V. DISCUSSION

### a) Developing a theory on dynamic transformations

The emphasis on construction using the Transform menu in GSP was shaped to facilitate the understanding of symmetries and strengthen the development of structures in the students' minds. The thought that the shapes have symmetries can lead students to dynamically reinvent new ways of constructing them through the dynamic geometry software.

The focus on transformations is in accordance to Coxford & Usiskin (1975), who report inter alia that, the use of different types of transformations in the curriculum simplifies the mathematical development (for example, the definitions of congruence and similarity cover all figures). Therefore, the proofs of many theorems are simpler and more accessible to all students. Furthermore, the authors argue that

transformations facilitate the understanding of mathematical concepts for students from different mathematical competence and prepare the ground for future processing concepts of algebra and analysis. Transformations used by the students in the DGS environment can be distinguished through the following:

- Theoretical and experimental dragging, as mentioned at the beginning of the article.
- Transformation elicited from the reflection, dilation, rotation, or translation of the object. Dragging on rotated (dilated, reflected, or translated) objects maintain the congruency and structural relationship between the elements of the construction.
- Transformations elicited from the utilization of the action buttons tools (for example, the hide/show action button, the link button, the movement button, or animation).
- Transformations elicited from the annotation of the dynamic diagram (for example, use of colors, formulations, and the trace tool). Moreover, the combination of transformations (e.g., the trace tool and dragging tool, the calculations and the dragging of the geometrical object's points).
- Transformations elicited from the application of the custom tools. The application of custom tools reorganizes the external representation. The application of a custom tool (or the repetition of the application of a custom tool) is accomplished in a sequence of steps directly perceived by the user. Consequently, custom tools operate as a referent point for organizing, pursuing, and retrieving information.
- Transformations elicited from the synthesis of the dynamic diagram.



- Combinations of transformations due to the synthesis of the software's interaction techniques (Sedig & Sumner, 2006).
- Complex transformations of the LVAR dynamic representations.

Therefore, *dynamic geometric transformations* are defined as the modifications of the diagram on screen that result in the modification in one or more included geometric objects. This could be an elicitation from the addition, cancelation of the diagram's elements that cause the rearrangement of the diagram, its anasynthesis, or even the modification of any object's

size or orientation. Moreover, it could be seen as we apply one or more interaction techniques, or their combination, on the diagram's objects. Transformations on prototype elements (e.g., points, line segments) led the students to (1) visualize the objects that were constructed in the first phase of the process and (2) perceive a few properties of the figure's symmetry initially at the visual level. It was observed that during the process the students connected, in their minds, representations that helped them to respond to the next level, according to the theory of van Hiele.

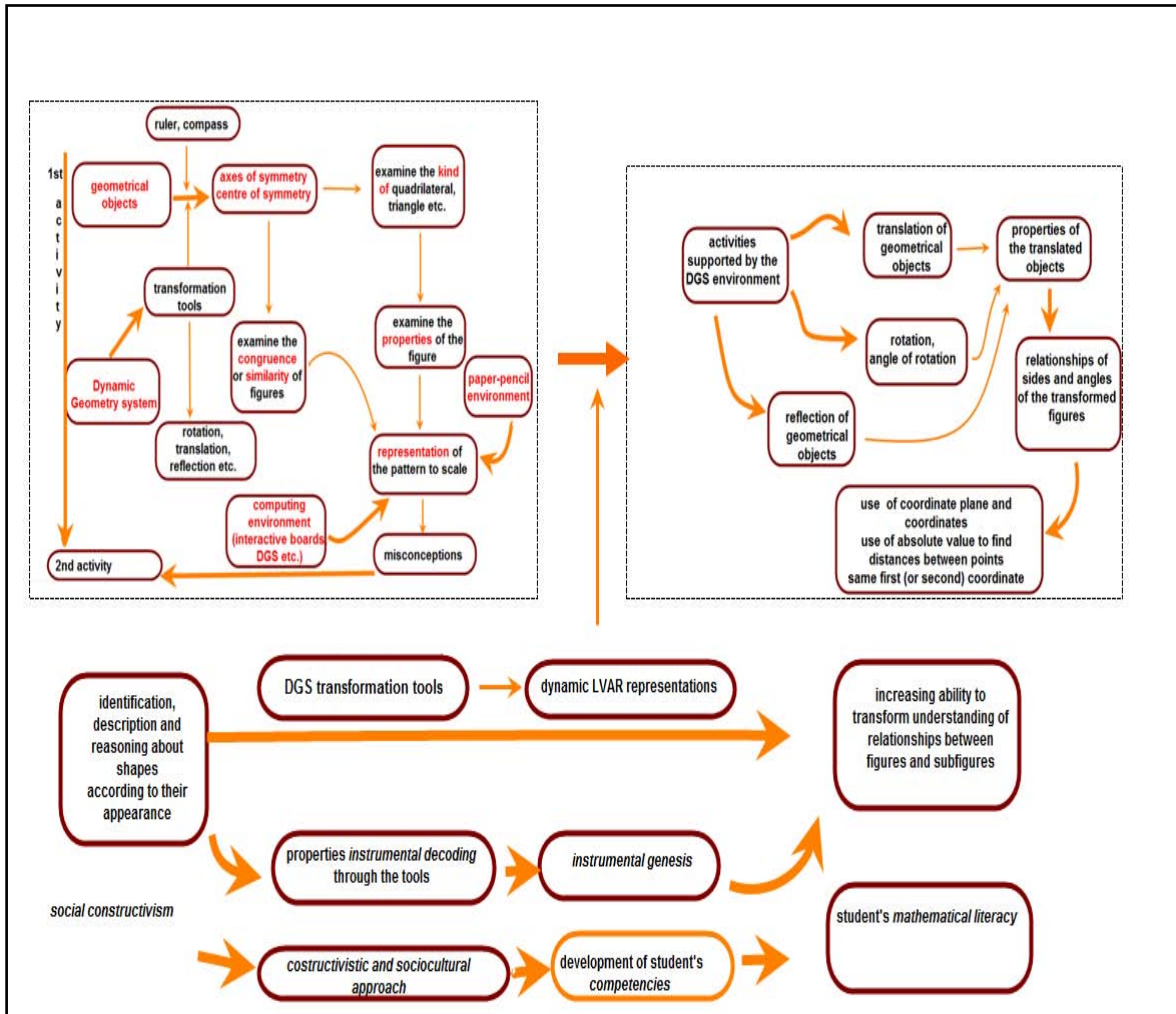


Figure 41 : A pseudo-Toulmin model explaining the transformations in the DGS environment

The dynamic manipulation of objects in software led the students to construct the properties of the shapes. The use of software transformation tools influenced the direction of their thinking, since their use allowed the properties of shapes to be analyzed and then synthesized into shapes. As a result, the construction and transformation of semi-preconstructed LVAR led the students to a theoretical way of thinking, and cognitive transformations through related cognitive connections.

"If we accept that mathematical growth coincides with constructing new mathematical reality, we may conceive mathematics education as supporting students in constructing new mathematical reality. This fits with Freudenthal's (1973) notion of "mathematics as a human activity". In his view students should be given the opportunity to reinvent mathematics. The challenge then is: How to make students invent what you want them to invent?" (Gravemeijer, 2004, p.3)

Many students do not have the ability to dynamically visualize and mentally manipulate geometric objects, which is important for solving problems in geometry. In that case they are not able to reflect on or to anticipate a possible solution to the problem. Therefore, geometric transformations in the software help the students to form an intermediate stage between the concrete and the abstract. They help them to instrumentally decode the mathematical symbols and to connect their use with the pre-existing knowledge. Then the interaction with the software incorporates the steps and the mental or cognitive actions that facilitate the understanding of the solution.

The use of the transformations in the DGS environment strongly influenced the formation of the 'dynamic' teaching cycle process which is described in the next section.

#### b) A 'dynamic' teaching cycle process through LVAR

The data presented here focused on teaching situations, including instructional units, classroom activities, and simulated or modeled problems in the DGS environment. In this section, I shall analyze my role as teacher, researcher, and instructor of the activities as it emerged from the teaching situations, as well as the students' role in the formation of a mathematical teaching cycle. The design or selection of teaching activities and problems that stimulate and excite mathematical reinvention (Freudenthal, 1973) on the part of students is a "challenge for the teacher, [who must] try to see the world through the eyes of the student." (Gravemeijer, 2004, p.8)

If the teaching and learning of concepts through the use of real problems in a DGS environment is compared with the traditional approach, we conclude that, "the modelling perspective [using a DGS environment] offers major advantages. The process of modelling constitutes the bridge between mathematics as a set of tools for describing aspects of the real world, on the one hand, and mathematics as the analysis of abstract structures, on the other" (Corte, Verschaffel & Greer, 2000, p.71). Moreover, the intrinsic design of dynamic representational systems has essential impacts on the mental representations of the student, that is, the ways in which students construct their personal representations of meaning during the activity, whether these representations are directed at an individual student or in the student's collaborative environment with others. Accordingly, the conclusions can be used to analyze the potential of these tools for mathematics teaching and learning, to design new tools, and to better understand the ways in which these tools can be (instrumentally) decoded by teachers and students to be transformed into theoretical knowledge built through mediation. As teachers (or teacher-researchers) design teaching concepts and ways of interacting with their students, they increasingly feel the need to understand

the minds of the students, looking for methods to lead their students to understand the concepts. Therefore, the determining factor is the teacher who decides on the objectives/aims of the teaching method and chooses the means for effective implementation of the objectives or of the educational process. The positive attitudes/behaviors of the teachers of mathematics with regard to mathematics, their positive position with regard to technology, and their interest in the students' understanding of the concepts, are the most important factors for the development of innovative applications in schools.

As a teacher-researcher, I know that the students encounter difficulties in order to understand the concepts in geometry. The connection between the represented and the representation can create conflicts to students because they are not able to control the information that comes from the outside world (Mesquita, 1998). The question is how we can overcome the cognitive obstacles they face and what are these teaching situations which can provide the scaffolding to the next van Hiele level.

The instructional units aimed to challenge students and "elicit, support and extend children's mathematical thinking, facilitating mathematical discussions, using the representations of concepts and encourage use of alternative solution methods" (Fuson et al., 2000, p. 277). Many times I tried to shift mentally from an observer's point of view to an actor's point of view (Cobb, Yackel & Wood, 1992 in Gravemeijer, 2004), and consider now the design of the activity of this regard. My approach was as follows:

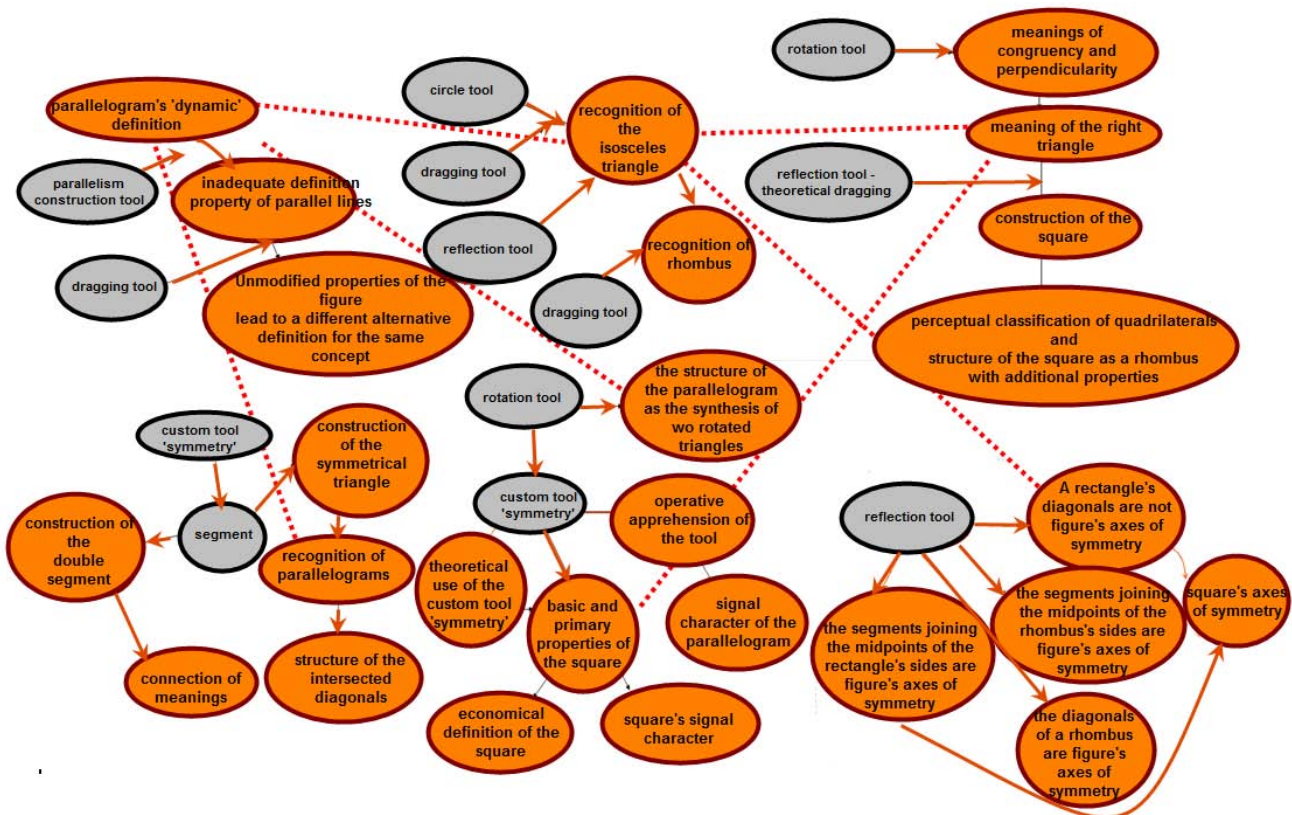


Figure 42 : A pseudo-Toulmin's model representing the instructional design through LVAR

*Situation 1 :* The highlighted idea in mathematics, which is 'symmetry,' is interdisciplinary, connected with art and culture. The aim is to 'see' mathematics in any context.

*Situation 2 :* The challenge was to connect the transformations in static and dynamic means conceptually and procedurally. Instrumental decoding of students' mathematical ideas played a major role for the overcoming of cognitive and instrumental obstacles.

*Situation 3:* This situation aims to accomplish the figure's symbol character. The grid in the DGS environment provides a challenge for the experimentation. The important points in this situation were the students' methods of dealing with the questions: "Under what conditions does the rhombus become a square?" or "What are the similarities and differences between a kite and a square?"

*Situation 4:* The motivation for this situation was that my students understand the parallelograms from their symmetry properties and, if they have a set of properties, to understand the kind of quadrilateral. This phase is very crucial for the students to acquire the ability to replace a figure with a set of properties that represent it and from these properties to construct the figure. In other words, the figure will acquire the signal character.

*Situation 5:* The recognition of differences and similarities between figures' symmetry properties demarcates the scope of this situation. The teaching

and didactic process must focus at the understanding of the structuring process and not the learning of ready-made structures.

*Situation 6:* The development of structures in students' minds has been achieved with the synthesis of a more complex construction. The situations aim to develop the abstraction. Pythagorean Theorem's reconfigurations have been used as a tool for the development of students' instrumental decoding of a complex figure's anasynthesis. The 6th situation led students to think about self similarity, which is not included in high school curriculum.

*Situation 7:* Self -similarity, Pythagorean Theorem and the midpoint theorem are the mathematical backgrounds of this situation. Here is explained the rationale in the design process and the importance of linking visual active representations and instrumental decoding.

The use of a computing environment such as dynamic geometry helps students to build 'a model of the meaning' (Thompson, 1987, p.85) and overcome the difficulties of translation between representations through the automatic translation or "dyna-linking" (Ainsworth, 1999, p. 133), since [they] "encode causal, functional, structural, and semantic properties and relationships of a represented world – either abstract or concrete" (Sedig & Sumner, 2006, p.2). The design and redesign of activities for the teaching and learning processes, with real problems or simulations of real-

world problems through LVAR in the dynamic geometry software, and the results obtained from the research data (Patsiomitou, 2012 a, b), suggest that a student develops his/her abstractive competency when his/her cognitive structures are linked through representations that the student develops during the learning process.

"Apart from the aspect of anticipating the mental activities of the students, a key element of the notion of a hypothetical learning trajectory is that the hypothetical character of the learning trajectory is taken seriously. The teacher has to investigate whether the thinking of the students actually evolves as conjectured, and he or she has to revise or adjust the learning trajectory on the basis of his or her findings. In relation to this, Simon (1995) speaks of a mathematical teaching cycle. In a similar manner, Freudenthal (1973) speaks of thought experiments that are followed by instructional experiments in a cyclic process of trial and adjustment. If we accept this image of the role of the teacher in instruction that aims at helping students to invent some (to them) new mathematics, we may ask ourselves, what type of support should be offered to teachers. Apparently, we will have to aim at developing means of support that teachers can use in constructing and revising hypothetical learning trajectories" (Gravemejer, 2004, p.9).

The whole action is an innovative production of a new approach to the educational process based on theoretical underpinning. This innovation is introduced for the first time in the school of established practice, and thus, proposes the redevelopment / redesign of the everyday teaching practice by using LVAR, with proper interventions in school curriculum. Specifically, linked representations that the student is able to construct (Patsiomitou, 2012a, b):

- When the student builds a representation (e.g., a rectangle) in order to create a robust construction externalizing his/her mental approach, using software interaction techniques by externalizing his/her mental approach or by transforming an external or internal representation to another representation in the same representational system or another one.
- When s/he gets feedback from the theoretical dragging to mentally link figures' properties so that, because of the addition of properties, subsequent representations stem from earlier ones.
- When s/he transforms representations so that the subsequent representations stem from previous ones due to the addition of properties.
- When s/he links mentally the developmental procedural aspects in a process of a dynamic reinvention
- When s/he reverses the procedure in order to create the same figure in a phase of the DHLP or between phases of the same DHLP.
- Adding to the initial [procedural] structure so that

the first component parts of a construction lead to a structure and to eventually becoming more and more complex,

- Linking the conceptual steps of the construction (p. 76).

Moreover, the procedures, due to their design, "prompted" the cooperation of students, contributed to the development of positive behavior, and strengthened the weak students to understand the concepts and procedures while interacting with their classmates. The process resulted in the cooperation of students with me which often "forgot" my role and "took" on the role of a student playing the 'game' to ask questions that some of my students did not have the courage to ask.

With regard to the problems, many teachers prefer algebra to geometry. The reasons are as follows: (a) the awareness of the risk of the student's failure or (b) the teacher's lack of confidence for their knowledge of the subject of geometry. How would the LVAR process (i.e. the utilization of LVAR concept for the construction of activities) change this weakness when students are able to process on official electronic platforms from the Ministry of Education? How will this affect the confidence of teachers who handle this platform for their students, giving feedback on their knowledge?

These questions should be discussed, as well as discussing who will educate the designers of these activities so that the material is consistent with the idea. On the other hand, it is obvious that there is possible misuse of the LVAR concept for the construction of activities by the way that every teacher thinks, which could lead to opposing results. It is therefore necessary to train the agents who will spread the LVAR idea, with consistent processes of meaning. Still, the implementation of the idea can be generalized and repeated in any group of students, at different times and in any thematic framework (e.g., the objects of physics or chemistry).

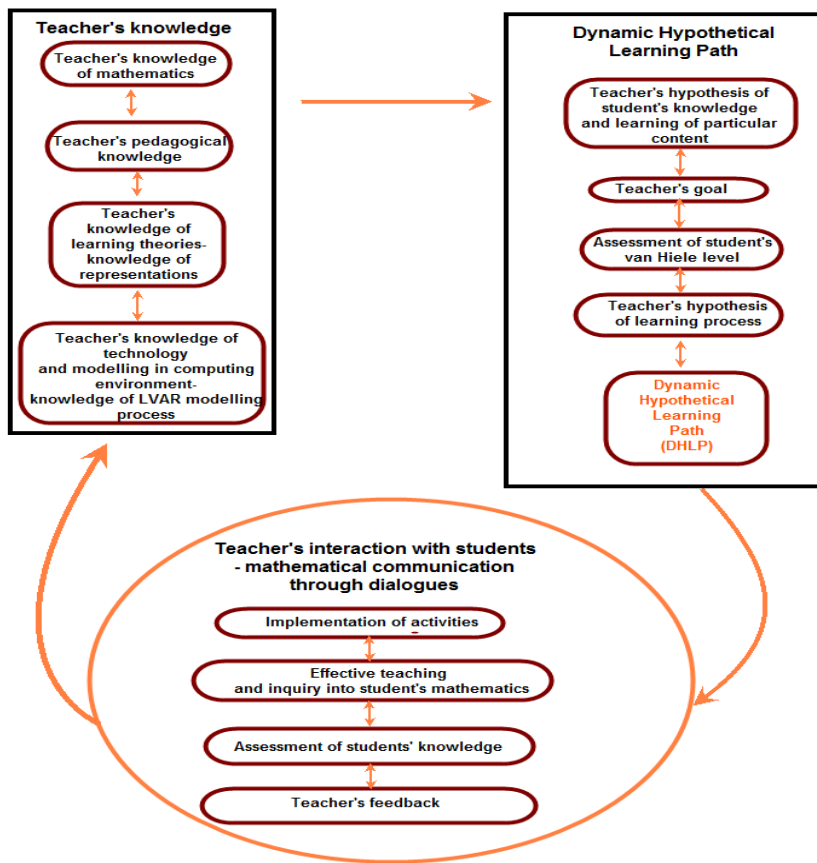


Figure 43 : An adaptation of Simon's (1995) Mathematics Teaching Cycle for the current study

The analysis of the teaching situations and the development of the Mathematics Teaching Cycle have led to an iterative diagram that is an adaptation of Simon's (1995, p.136) work, taking into account also the work of McGraw (2002, p.10). The Mathematics Teaching Cycle portrays the relationship between the following areas of knowledge (Simon, 1995): "the teacher's knowledge of mathematics mathematics and his hypotheses about the students' understandings, several areas of teacher knowledge come into play, including the teacher's theories about mathematics teaching and learning; knowledge of learning with respect to the particular mathematical content; and knowledge of mathematical representations, materials, and activities" (p. 133).

What has been examined is the use of technology in the teaching cycle which plays an important role in the development of discussions, as well as students' vH level. The diagram aims to include the incorporation of technology practices in class. The teacher's interaction with students and the mathematical communication through dialogues is accomplished in sequential situations: the implementation of activities, effective teaching and inquiry into students' mathematics, the assignment of students' knowledge, all of which leads to the teacher's feedback. These processes go on continually and can suggest

adaptations in various domains of a teacher's knowledge, including in the following areas: mathematics, pedagogy, representations, technology, and modeling through LVAR representations. The whole process leads to a modification of the hypothetical learning path that includes a continuous interaction between the teacher's knowledge of particular content, the teacher's goal, and assessment of the students' vH levels.

## VI. CONCLUSIONS

The modeling of a problem in the dynamic environment can 'carry' any [mathematical] object to the classroom in two ways: through the use of digital images or through the use of their simulations. On the other hand, a technological tool is important as the design of artifacts can be generalized and replicated in any group of students, at different times and in any thematic framework (e.g., science, geography). Therefore, referring to LVAR is concluded in the following (Patsiomitou, 2012a, p 498):

- How could this affect the students' understanding of the utilization of LVAR in the teaching and learning of other disciplines (e.g., physics or ancient Greek and history)? [or] Would students understand the obscure points of other disciplines, because of the

interaction with the [appropriate] dynamic LVAR representations?

- Can the students develop their linking of the conceptual and procedural representations of these objects?

On the other hand, new cognitive tools are not included [or included in a very slow way] for the teaching of concepts. It is particularly important for the 'movement' of a process by applying innovative practices to change the negative views that a large portion of teachers have regarding technology. This seems to focus on a lack of knowledge because of the phobias surrounding technological tools in the mathematics classroom, leading to an adherence to traditional teaching methods.

In general, the whole issue has to do with the way we perceive the world, the natural objects (unconscious), how we compare them mentally (consciously) with theoretical constructs of geometry in order to represent them and how we instrumentally decode them using technology. Finally, it is important to continue teaching and research concepts in this vital field, through activities that involve children in the learning process, so using linked visual representations they will learn how to develop, interpret, and make sense of geometric concepts. This argument recognizes and underlines the force of Kant's argument (1929, "Critique of Pure Reason") that:

There can be no doubt that all our knowledge begins with experience. For how should our faculty of knowledge be awakened into action did not objects affecting our senses partly of themselves *produce representations*, partly arouse the activity of our understanding to *compare these representations*, and, by *combining or separating them*, work up the raw material of the sensible impressions into that knowledge of objects which is entitled experience? [Because] "Understanding is the faculty of knowledge and [...] *knowledge consists in the determinate relation of given representations to an object*".

## REFERENCES RÉFÉRENCES REFERENCIAS

- Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education*, 33, 131-152.
- Ainsworth, S. E. (2006). DeFT: A conceptual framework for learning with multiple representations. *Learning and Instruction*, 16(3), 183-198.
- Almeqdadi, F. (2000). *The effect of using the geometer's sketchpad (GSP) on Jordanian students' understanding of geometrical concepts*. Jordan: Yarmouk University.
- Bartolini Bussi M. G. (1996). Mathematical Discussion and Perspective Drawing in Primary School, *Educational Studies in Mathematics*, 31, 11-41.
- Battista, M. T. (2007). The development of geometric and spatial thinking. In Lester, F. (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 843-908). NCTM. Reston, VA: National Council of Teachers of Mathematics.
- Battista, M. T. (2011). Conceptualizations and issues related to learning progressions, learning trajectories, and levels of sophistication. *The Mathematics Enthusiast*, 8, 507-570. [http://www.math.umd.edu/TMME/vol8no3/Battista\\_TME\\_2011\\_article4\\_pp.507\\_570.pdf](http://www.math.umd.edu/TMME/vol8no3/Battista_TME_2011_article4_pp.507_570.pdf)
- Bauersfeld, H. (1992). Integrating Theories for Mathematics Education. *For the Learning of Mathematics*, 12(2), 19-28.
- Bruner, J. S. (1961). The act of discovery. *Harvard Educational Review*, 31, 21-32.
- Bruner, J. S. (1966). *Toward a theory of instruction*, Cambridge, Mass.: Belkapp Press.
- Bruner, J. S. (1986) *Actual Minds, Possible Worlds*. Cambridge, Mass: Harvard University Press.
- Burkhardt, H. (1981). *The real world and mathematics*. Glasgow, UK: Blackie Cambridge, Mass.: Harvard University Press, 1986
- Carpenter, T.P., Fennema, E., Peterson, P.L., Chiang, C., & Loef, M. (1989). Using knowledge of children's mathematics thinking in classroom teaching: An experimental study. *American Educational Research Journal*, 26(4), 499-531.
- Carpenter, T. P., Fennema, E., Franke, M., Levi, L., & Empson, S. B. (1999). *Children's mathematics: Cognitively Guided Instruction*. Portsmouth, NH: Heinemann.
- Choppin, J. M. (1994). Spiral through recursion. *Mathematics Teacher*, 87, 504-508.
- Clement, D.H. & Battista, M.T. (1990). Research into practise: Constructivist learning and teaching. *Arithmetic Teacher*, 38 (1), 34-35.
- Cobb, P., & Steffe, L. P. (1983). The constructivist researcher as teacher and model builder. *Journal for Research in Mathematics Education*, 14(2), 83-94.
- Cobb, P., & Bauersfeld, H. (Eds.). (1995). The emergence of mathematical meaning: Interaction in classroom cultures. Hillsdale, NJ: Erlbaum
- Cobb, P., Yackel, E., & Wood, T. (1992). A constructivist alternative to the representational view of mind in mathematics education. *Journal for Research in Mathematics Education*, 23(1) 2-33.
- Corcoran, T., Mosher, F.A., & Rogat, A. 2009. Learning Progressions in Science: An Evidence-based Approach to Reform. New York, NY: Center on Continuous Instructional Improvement, Teachers

- College, Columbia University. Retrieved from [http://www.cpre.org/images/stories/cpre\\_pdfs/lp\\_science\\_rr63.pdf](http://www.cpre.org/images/stories/cpre_pdfs/lp_science_rr63.pdf)
20. Coxford, A. F., Usiskin Z. P.(1975): *Geometry: A Transformation Approach*. Laidlaw Brothers, Publishers.
  21. De Corte, E., Verschaffel, L., & Greer, B. (2000). Connecting mathematics problem solving to the real world. In: Proceedings of the International Conference on Mathematics Education into the 21st Century: Mathematics for living (pp. 66-73). Amman, Jordan: The National Center for Human Resource Development.
  22. Drijvers, P., & Trouche, L. (2008). From artifacts to instruments: A theoretical framework behind the orchestra metaphor. In G. W. Blume & M. K. Heid (Eds.), *Research on technology and the teaching and learning of mathematics: Cases and perspectives, 2*, 363-392. Retrieved from [https://www.academia.edu/4969563/Drijvers\\_P\\_and\\_Trouche\\_L\\_2008\\_From\\_artifacts\\_to\\_instruments\\_a\\_theoretical\\_framework\\_behind\\_the\\_orchestra\\_metaphor](https://www.academia.edu/4969563/Drijvers_P_and_Trouche_L_2008_From_artifacts_to_instruments_a_theoretical_framework_behind_the_orchestra_metaphor).
  23. Docq, F., & Daele, A. (2001). Uses of ICT tools for CSDL: how do students make as their's own the designed environment? In *Proceedings of the First European Conference on Computer-Supported Collaborative Learning* (pp. 197-204). Maastricht, Netherlands. Retrieved from [www.ipm.ucl.ac.be/articlesetsupportsIPM/eurocsclDocq.pdf](http://www.ipm.ucl.ac.be/articlesetsupportsIPM/eurocsclDocq.pdf)
  24. Dougiamas, M. An exploration of the use of an Open Source software called Moodle to support a social constructionist epistemology of teaching and learning within Internet-based communities of reflective inquiry. *Science and Mathematics Education Centre, Curtin University of Technology*. PhD thesis. (<http://dougiamas.com/thesis/>)
  25. Duval, R. (1995). *Sémiosis et pensée humaine*. Berne: Peter Lang
  26. Duschl, A., Schweingruber, H., & Shouse A. (2007). *Taking science to school: Learning and teaching science in grades K-8*. Washington, DC: National Academies Press.
  27. Ferrara, F., Pratt, D., & Robutti, O. (2006). The role and uses of technologies for the teaching of algebra and calculus: Ideas discussed at PME over the last 30 years. In A. Gutierrez & P. Boero (Eds.), *Handbook of research on the psychology of mathematics education: Past, present and future*. (pp. 237-273). Rotterdam, The Netherlands: Sense Fractovia. <http://www.fractovia.org/art/people/mandeIbrot.html>.
  28. Freudenthal, H. (1973). *Mathematics as an Educational Task*. Dordrecht: Reidel, The Netherlands
  29. Freudenthal, H. (1983). *Didactical phenomenology of mathematical structures* (p. 595). Dordrecht, The Netherlands: D. Reidel.
  30. Freudenthal (1991), *Revisiting Mathematics Education: China Lectures*. Dordrecht: Kluwer Academic Publishers.
  31. Fuson, K. C., Carroll, W. M., & Drueck, J. V. (2000). Achievement results for second and third graders using the Standards-based curriculum. *Everyday Mathematics. Journal for Research in Mathematics Education, 31*, 277-295.
  32. Fuys, D., Geddes, D., & Tischler, R. (Eds.). (1984). *English translation of selected writings of Dina van Hiele-Geldof and Pierre M. van Hiele*. Brooklyn: Brooklyn College. (ERIC Document Reproduction Service No. ED 287 697).
  33. Fuys, D., Geddes, D., & Tischler, R. (1988). The Van Hiele model of thinking in geometry among adolescents. *Journal for Research in Mathematics Education* : Monograph Number 3.
  34. Giraldo, V., Belfort, E., & Carvalho, L. M. (2004). Descriptions and conflicts in dynamic geometry. In M. J. Heines & A. B. Fuglestad (Eds.), *Proceedings of the 28th PME International Conference, 2*, 455-462.
  35. Glasersfeld, E. v. (1995). *Radical constructivism: A way of knowing and learning*. London: Falmer Press
  36. González G., Herbst P. (2009) Students' conceptions of congruency through the use of dynamic geometry software. *International Journal of Computers for Mathematical Learning, 14* (2) 153-182.
  37. Goldenberg, P. (1999), Principles, Art, and Craft in Curriculum Design: The Case of Connected Geometry, *International Journal of Computers for Mathematical Learning 4*: Kluwer
  38. Goldin, G. (2003). Representation in school mathematics: A unifying research perspective. In J. Kilpatrick, Martin, W., Schifter, D., & National Council of Teachers of Mathematics (Ed.). *A research companion to Principles and standards for school mathematics*, 17, 1-4. Reston: National Council of Teachers of Mathematics.
  39. Goos, M. (2004). Learning mathematics in a classroom community of inquiry. *Journal for Research in Mathematics Education, 35*, 258-291.
  40. Goos, M., Galbraith, P., & Renshaw, P. (2002). Socially mediated metacognition: creating

collaborative zones of proximal development in small group problem solving. *Educational Studies in Mathematics*, 49, 193–223

41. Goos, M., Galbraith, P., Renshaw, P., & Geiger, V. (2003) Perspectives on technology mediated learning in secondary school mathematics classrooms. *Journal of Mathematical Behavior*, 22, 73-89.
42. Govender, R. & De Villiers, M. (2004). A dynamic approach to quadrilateral definitions. *Pythagoras* 58, 34-45.
43. Govender, R., & de Villiers, M (2002). Constructive Evaluations of Definitions in a Sketchpad Context. Paper presented at the Association for Mathematics Education of South Africa National conference, University of Natal, Durban, July 2002. [web.co.za/residents.profmd/rajen.pdf](http://web.co.za/residents.profmd/rajen.pdf).
44. Gravemeijer, K. P. E. (2004). *Creating Opportunities for Students to Reinvent Mathematics*. Paper presented at ICME 10, Copenhagen, Denmark. July 4-11. Retrieved from <http://www.staff.science.uu.nl/~savel101/edsci10/literature/gravemeijer1994.pdf>
45. Gutierrez, A. & Jaime, A. (1998). On the assessment of the Van Hiele levels of reasoning, *Focus on Learning Problems in Mathematics* 20 (2/3), 27-46. Retrieved from [www.uv.es/Angel.Gutierrez](http://www.uv.es/Angel.Gutierrez)
46. Hadas, N., Hershkowitz, R., & Schwarz, B. B. (2000). The role of contradiction and uncertainty in promoting the need to prove in dynamic geometry environments. *Educational Studies in Mathematics*, 44(1/2), 127-150.
47. Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of writing processes. In L. W. Gregg & E. R. Sternberg (Eds.), *Cognitive processes in writing*. Hillsdale NJ: Erlbaum.
48. Hershkowitz, R. (1990) 'Psychological aspects of learning geometry', in P. Nesher and J. Kilpatrick (eds), *Mathematics and Cognition*, Cambridge University Press, Cambridge
49. Hoffer, A. (1981). Geometry is more than proof. *Mathematics Teacher*, 74, 11-18
50. Hollebrands, K. (2007). The role of a dynamic software program for geometry in the strategies high school mathematics students employ. *Journal for Research in Mathematics Education*, 38(2), 164-192.
51. Hollebrands, K.F. & Smith, R.C. (2009). Using dynamic geometry software to teach secondary school geometry: Implications from research. In T. Craine (Ed.) *Understanding Geometry for a Changing World: 2009 Yearbook of the NCTM*. (pp. TBD). Reston: VA: National Council of Teachers of Mathematics.
52. Jackiw, N. (1991). *The Geometer's Sketchpad* [Computer Software]. Berkeley, CA: Key Curriculum Press
53. Jackiw, N. (2006) Mathematics and the Aesthetic: New Approaches to an Ancient Affinity. *Mechanism and Magic in the Psychology of Dynamic Geometry*. Heidelberg: Springer Verlag.
54. Jackiw, N, Sinclair N. (2004) Iteration and Dynamic Geometry: Beyond Common Fractals with The Geometer's Sketchpad *NCTM Philadelphia Session 142*
55. Jaworski, B. (2003). Inquiry as a pervasive pedagogic process in mathematics education development. *Proceedings of the Third Conference of the European Society for Research in Mathematics Education*. Bellaria, Italy. Retrieved from <http://www.dm.unipi.it/~didattica/CERME3>
56. Johnson-Laird, P. N. (1983). *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Cambridge: Cambridge University Press; Cambridge, MA: Harvard University Press.
57. Kadunz, G. & Sträßer, R. (2004). Image - metaphor - diagram: visualisation in learning mathematics *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education*, Vol 4, pp 241–248.
58. Kant, I. (1929). *Critique of Pure Reason*. Translated by Norman Kemp Smith. Basingstoke: Hants: Palgrave.
59. Kaput, J. (1991). Notations and representations as mediators of constructive processes. In E. Glasersfeld (Ed.), *Radical Constructivism in Mathematics Education* (pp. 53-74). Netherlands: Kluwer Academic Publishers.
60. Kaput, J. (1999). Representations, inscriptions, descriptions and learning: A kaleidoscope of windows. *Journal of Mathematical Behavior*, 17(2), 256-281.
61. Kaput, J. J. (1992). Technology and Mathematics Education. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 515-556). New York: Macmillan
62. Kline, M. (1990). *Mathematics in Western Culture*. Penguin, London.
63. Krajcik, J, Shin, N., Stevens, S. Y., & Short, H. (2009) Using Learning Progressions to Inform the Design of Coherent Science Curriculum Materials. Paper presented at the *Annual Meeting of the American Education Research Association*, San Diego, CA. Retrieved from [http://www.umich.edu/~hiceweb/PDFs/2009/AERA\\_LP\\_UMichigan.pdf](http://www.umich.edu/~hiceweb/PDFs/2009/AERA_LP_UMichigan.pdf)



64. Krajcik, J, Shin, N., Stevens, S. Y., & Short, H. (2009) Using Learning Progressions to Inform the Design of Coherent Science Curriculum Materials. Paper presented at the Annual Meeting of the American Education Research Association, San Diego, CA.
65. Leung, A & Or C.M (2007) From construction to proof: explanations in dynamic geometry environment. In Woo, J. H., Lew, H. C., Park, K. S. & Seo, D. Y. (Eds.). Proceedings of the 31st Conference of the International Group for the Psychology of Mathematics Education, Vol. 3, pp. 177-184. Seoul: PME.
66. Long, M. C. (2011) Mathematical, cognitive and didactic elements of the multiplicative conceptual field investigated within a Rasch assessment and measurement framework. Doctoral dissertation. University of Cape Town. Retrieved from [http://web.up.ac.za/sitefiles/file/43/314/Long\\_M\\_C\\_\(2011\).The\\_multiplicative\\_conceptual\\_field\\_investigated\\_within\\_a\\_Rasch\\_measurement\\_framework\\_PDF](http://web.up.ac.za/sitefiles/file/43/314/Long_M_C_(2011).The_multiplicative_conceptual_field_investigated_within_a_Rasch_measurement_framework_PDF)
67. Mariotti M.-A. (2000). Introduction to proof: The mediation of a dynamic software environment. *Educational Studies in Mathematics*, 44(1-3), 25-53.
68. Martinez, Juan. "BENOÎT MANDELBROT." Third. Apex.to.fractovia. 02 19 2003. Third. Apex. to. Fractovia. Retrieved from <http://www.fractovia.org/art/people/mandelbrot.html>.
69. McGraw, R. H. (2002). *Facilitating whole-class discussion in secondary mathematics classrooms*. Unpublished doctoral dissertation, Indiana University, Bloomington, IN.
70. Mesquita, A. (1998). On conceptual obstacles linked with external representation in geometry. *Journal of Mathematical Behavior*, 17(2), 183-195.
71. Mogetta, C., Olivero, F. & Jones, K. (1999) Providing the Motivation to Prove in a Dynamic Geometry Environment. Proc. of the British Society for Research into Learning Mathematics, St Martin's University College, Lancaster, 1999, pp91-6.
72. National Research Council (1996). *National Science Education Standards*. Washington, D.C: National Academy Press.
73. National Research Council. (2001). *Knowing what students know: The science and design of educational assessment*. Committee on the Foundations of Assessment. J. Pelligrino, N. Chudowsky, & R. Glaser (Eds.), Board on Testing and Assessment, Center for Education, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.
74. National Research Council. (July 12-August 2, 2010). *A framework for science education: Preliminary public draft*. Washington, DC: National Academy Press
75. Neubrand, M. (Ed.) (2004). *Mathematische Kompetenzen von Schülerinnen und Schülern in Deutschland: Vertiefende Analysen im Rahmen von PISA-2000*. Wiesbaden: VS – Verlag für Sozialwissenschaften.
76. Neubrand, M., Biehler, R., Cohors-Fresenborg, E., Flede, L., Knoche, N., Lind, D., Loding, W., Moller, G., Wynands, A. (2001). Grundlager der Ergänzung des Internationalen PISA-Mathematik-Tests in der Deutschen Zusatzerhebung. Retrieved from <http://link.springer.com/article/10.1007%2FBF02652739#page-2>
77. Niss, M. (1999), Kompetencer og Uddannelsesbeskrivelse (Competencies and subject description). *Uddanneise*, 9, 21-29.
78. Niss, M. (2003). Mathematical competencies and the learning of mathematics: The Danish KOM project. Third Mediterranean Conference on Mathematics Education, Athens (pp. 115–124).
79. OECD (2006), *Assessing Scientific, Reading and Mathematical Literacy – A Framework for PISA 2006*, Paris: OECD.
80. OECD (2010). *PISA 2012 Mathematics Framework .Paris: OECD Publications*. Retrieved from <http://www.oecd.org/dataoecd/8/38/46961598.pdf>
81. Olkun, S., Sinoplu, N. B., & Deryakulu, D. (2005). Geometric exploration with dynamic geometry applications based on van Hiele levels. *International Journal for Mathematics Teaching and Learning*. Retrieved from <http://www.cimt.plymouth.ac.uk/journal/olkun.pdf>
82. Parzysz, B. (1988). Knowing versus seeing: problems of the plane representation of space geometry figures. *Educational Studies in Mathematics*, 19, 79–92.
83. Patsiomitou, S. (2005) Fractals as a context of comprehension of the meaning of sequence in a Dynamic Geometry environment. Master Thesis. Department of Mathematics of the National and Kapodistrian University of Athens (in cooperation with the University of Cyprus).
84. Patsiomitou, S. (2007) Fractals as a context of comprehension of the meanings of the sequence and the limit in a Dynamic Software environment. *Electronic Proceedings of the 8th International Conference on Technology in Mathematics Teaching (ICTMT8)* in Hradec Králové (E. Milková,

- Pavel Prazák, eds.), University of Hradec Králové, ISBN 397-80-7041-285-5 (cd-rom).
85. Patsiomitou, S. (2008a). The development of students' geometrical thinking through transformational processes and interaction techniques in a dynamic geometry environment. *Issues in Informing Science and Information Technology Journal*, 5, 353-393. <http://iisit.org/IssuesVol5.htm>
  86. Patsiomitou, S. (2008b) Linking Visual Active Representations and the van Hiele model of geometrical thinking. *Proceedings of the 13<sup>th</sup> Asian Conference in Technology in Mathematics*. pp. 163-178. Bangkok, Thailand: Suan Shunanda Rajabhat University. <http://atcm.mathandtech.org/EP2008/pages/regular.htm>
  87. Patsiomitou, S., (2008c) Do geometrical constructions affect students algebraic expressions? *Proceedings of the 13<sup>th</sup> Asian Conference in Technology in Mathematics*. pp. 193-202 Bangkok, Thailand: Suan Shunanda Rajabhat University. <http://atcm.mathandtech.org/EP2008/pages/regular.htm>
  88. Patsiomitou, S., (2008d) Custom tools and the iteration process as the referent point for the construction of meanings in a DGS environment. *Proceedings of the 13<sup>th</sup> Asian Conference in Technology in Mathematics*. pp.179-192 Bangkok, Thailand: Suan Shunanda Rajabhat University. <http://atcm.mathandtech.org/EP2008/pages/regular.html>
  89. Patsiomitou, S. (2009) The Impact of Structural Algebraic Units on Students' Algebraic Thinking in a DGS Environment. *Electronic Journal of Mathematics and Technology (eJMT)*, 3(3), 243-260.
  90. Patsiomitou, S. (2010a). Building LVAR (Linking Visual Active Representations) modes in a DGS environment. *Electronic Journal of Mathematics and Technology*, 4(1), 1-25.
  91. Patsiomitou, S. (2010b). Learning mathematics with the Geometer's Sketchpad v4, Athens: Kleidarithmos (in Greek)
  92. Patsiomitou, S. (2011). Theoretical dragging: A non-linguistic warrant leading to dynamic propositions. *Proceedings of the 35<sup>th</sup> Conference of the International Group for the Psychology of Mathematics Education*, Vol. 1, pp. 361-368. Ankara, Turkey: PME
  93. Patsiomitou, S. (2012a). The development of students' geometrical thinking through transformational processes and interaction techniques in a DGS environment. Linking Visual Active Representations. Unpublished Ph.D. Thesis. University of Ioannina, Ioannina (December, 2012)
  94. Patsiomitou, S. (2012b) A Linking Visual Representation DHLP for student's cognitive development. *Global Journal Of Computer Science and Technology*. 12 (6), 53-82
  95. Patsiomitou, S. (2013). Students learning paths as 'dynamic encephalographs' of their cognitive development. *International journal of computers & technology* [Online], 4(3), 802-806. (18 April 2013)
  96. Patsiomitou, S. & Koleza, E. (2008). Developing students' geometrical thinking through linking representations in a dynamic geometry environment. In O. Figueras, & Sepúlveda, A. (Eds.). *Proceedings of the Joint Meeting of the 32<sup>nd</sup> Conference of the International Group for the Psychology of Mathematics Education* (pp. 89-96). Morelia, Michoacán, México: PME
  97. Patsiomitou, S. and Emvalotis A. (2009a). Developing geometric thinking skills through dynamic diagram transformations. *Proceedings of MEDCONF 2009, The Sixth Mediterranean Conference on Mathematics Education*. pp.249-258. Plovdiv, Bulgaria.
  98. Patsiomitou, S. & Emvalotis A. (2009b). Does the Building and transforming on LVAR modes impact students' way of thinking? In Tzekaki, M., Kaldrimidou, M. & Sakonidis, C. (Eds.). *Proceedings of the 33<sup>rd</sup> Conference of the International Group for the Psychology of Mathematics Education*, Vol. 4, pp. 337-344. Thessaloniki, Greece: PME.
  99. Patsiomitou, S. & Emvalotis A. (2009c) 'Economy' and 'Catachrèse' in the use of custom tools in a Dynamic geometry problem-solving process. *Electronic Proceedings of the 9<sup>th</sup> International Conference on Technology in Mathematics Teaching (ICTMT9)* in Metz, France.
  100. Patsiomitou, S. & Emvalotis A. (2009d) Composing and testing a DG research-based curriculum designed to develop students' geometrical thinking. Paper at the annual European Conference on Educational Research (ECER) Vienna. <http://www.eera-ecer.eu/ecer-programmes-and-presentations/conference/ecer-2009/contribution/1961>
  101. Patsiomitou, S. & Koleza, E. (2009). The development of students' geometrical thinking through Linking Visual Active Representations. In M. Kourkoulos & Tzanakis, C. (Eds.). *Proceedings of the 5<sup>th</sup> International Colloquium on the Didactics of Mathematics* (Vol. II), pp.157-171. Rethymnon, Greece.

102. Patsiomitou, S & Emvalotis, A. (2010a). The development of students' geometrical thinking through a DGS reinvention process. In M. Pinto, & Kawasaki, T. (Eds), *Proceedings of the 34th Conference of the International Group for the Psychology of Mathematics Education*, Vol.4 (pp. 33-40), Belo Horizonte. Brazil: PME
103. Patsiomitou, S. & Emvalotis A. (2010b). Students' movement through van Hiele levels in a dynamic geometry *guided reinvention* process. *Journal of Mathematics and Technology* (JMT), (3), 18-48.
104. Patsiomitou, S., Barkatsas, A., Emvalotis, A. (2010). Secondary students' "dynamic reinvention of geometric proof" through the utilization of linking visual active representations, *Journal of Mathematics and Technology*, 5, 43-56.
105. Patton, M. (1990). *Qualitative evaluation methods (2nd ed.)*. Thousand Oaks, CA: Sage.
106. Pellegrino, J.W., Chudowsky, N., & Glaser, R. (Eds.). (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press.
107. Piaget, J. (1937/1971). *The construction of reality in the child* (M. Cook, Trans.). New York: Basic Books.
108. Peirce, C. S. (1998, c. 1903d). The three normative sciences. In N. Houser & C. Kloesel, (Eds.), *The Essential Peirce: Selected philosophical writings*, Volume 2, 196-207. Bloomington: Indiana University Press.
109. Pierce R, Stacey K. (2009). Using Dynamic Geometry to Bring the Real World into the Classroom. Retrieved from [http://www.edfac.unimelb.edu.au/sme/research/Pierce\\_Stacey\\_GGB.pdf](http://www.edfac.unimelb.edu.au/sme/research/Pierce_Stacey_GGB.pdf).
110. Schumann, H. (2004): Reconstructive Modelling with Dynamic Geometry Systems. In: *EduMath* 19, (12), 3 – 21.
111. Schumann, H. & Green, D. (1994). *Discovering geometry with a computer – using Cabri Geometre*. Sweden: Studentlitteratur. Lund
112. Sedig, K., & Sumner, M. (2006). Characterizing interaction with visual mathematical representations. *International Journal of Computers for Mathematical Learning*, 11, 1-55. New York: Springer.
113. Sfard, A. (2001). Learning mathematics as developing a discourse. In R. Speiser, C. Maher, C. Walter (Eds), *Proceedings of 21st Conference of PME-NA* (pp. 23-44). Columbus, Ohio: Clearing House for Science, mathematics, and Environmental Education. Retrieved from [http://mathcenter-k6.haifa.ac.il/articles\(pdf\)/sfard.pdf](http://mathcenter-k6.haifa.ac.il/articles(pdf)/sfard.pdf)
114. Simon, M. A. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for Research in Mathematics Education*, 26 (2), 114-145
115. Simon, M.A. & Shifter, D. (1991). Towards a constructivist perspective: An intervention study of mathematics teacher development. *Educational Studies in Mathematics*, 22(4), 309-331.
116. Sinclair N., & Jackiw, N. (2007). Modeling practices with The Geometer's Sketchpad. *Proceedings of the ICTMA-13*. Bloomington, IL: Indiana University
117. Sinclair, M.P. (2001). Supporting student efforts to learn with understanding: an investigation of the use of Javasketchpad sketches in the secondary school geometry classroom. *Unpublished doctoral dissertation*, University of Toronto, Graduate Department of Education.
118. Skemp, R. (1986). *The psychology of learning mathematics (2nd ed.)* Harmondsworth, UK: Penguin
119. Smith, C. L., Wiser, M., Anderson, C. W., & Krajcik, J. (2006). Implications of research on children's learning for standards and assessment: A proposed learning progression for matter and the atomic molecular theory. *Measurement: Interdisciplinary Research and Perspectives*, 14, 1-98.
120. Smith, P. L., & Wedman, J. F. (1988). Read-think-aloud protocols: A new data-source for formative evaluation. *Performance Improvement Quarterly*, 1(2), 13-22.
121. Stacey, K. (2012). The international assessment of mathematical literacy: PISA 2012 framework and items. *Paper presented at the 12th International Congress on Mathematical Education*, COEX, Seoul, Korea, July 8-15. Retrieved from [http://www.icme12.org/upload/submission/2001\\_F.pdf](http://www.icme12.org/upload/submission/2001_F.pdf)
122. Steffe, L. P., & Gale, J. E. (1995). *Constructivism in education*. Hillsdale, NJ: Lawrence Erlbaum.
123. Steketee, S. (2002) Iteration through the math curriculum :Sketchpad 4 does it again and again *NCTM Annual Meeting Session 491*
124. Stevens, S., Shin, N, & Krajcik, J. (2009). Towards a model for the development of an empirically tested learning progression. Paper presented at the Learning Progressions in Science (LeaPS) Conference, Iowa City, IA. Retrieved from <http://www.education.uiowa.edu/projects/leaps/proceedings/>
125. Thompson, P. (1987). Mathematical microworlds and intelligent computer-assisted instruction. In G. Kearsley (Ed.), *Artificial intelligence and instruction: Applications and methods* (pp. 83-109). Reading: Addison
126. Toulmin, S.E. (1958). *The uses of argument*. Cambridge: Cambridge University Press.
127. Trouche, L. (2004). Managing the complexity of the human/machine interaction in computerized learning environments: Guiding students' command process through instrumental orchestrations.

- International Journal of Computers for Mathematical Learning* 9, 281-307. Kluwer Academic Publishers.
128. Usiskin, Z. (1982). *Van Hiele levels and achievement in secondary school geometry*. (Final report of the Cognitive Development and Achievement in Secondary School Geometry Project). Chicago: University of Chicago. (ERIC Document Reproduction Service, No. ED 220 288).
129. Üstün, I & Ubuz, B. (2004) Student's Development of Geometrical Concepts Through a Dynamic Learning Environment, TSG 16: Visualisation in the teaching and learning of mathematics. Retrieved from [www.icme-organisers.dk/tsg16/papers/Ubuz.TSG16.pdf](http://www.icme-organisers.dk/tsg16/papers/Ubuz.TSG16.pdf)
130. van Hiele, P. M. (1986). *Structure and insight: A theory of mathematics education*. Orlando, Florida: Academic Press.
131. Vergnaud, G. (1987). Conclusion. In C. Janvier (Ed.), *Problems of Representations in the Teaching and Learning of Mathematics* (pp. 227-232). Hillsdale: Lawrence Erlbaum Associates.
132. Vergnaud, G. (1988). Multiplicative structures. In J. Hiebert & M. Behr (Eds.), *Number concepts and operations in the middle grades*. Hillsdale, NJ: National Council of Teachers of Mathematics.
133. Vergnaud, G. (1998). Comprehensive theory of representation for mathematics education. *Journal of Mathematics Behavior*, 17(2), 167-181.
134. Vinner, S. (1983). Concept Definition Concept Image and the Notion of Function. *International Journal for Mathematics Education in Science and Technology*, 14, (3), 293-305.
135. Vygotsky, L.S. (1987) Thinking and speech. In R.W. Rieber and A.S. Carton (Eds.), *The collected works of L.S. Vygotsky, Volume 1: Problems of general psychology*, (Trans. N.Minick). New York: Plenum.
136. White, P., & Mitchelmore, M. C. (2010) Teaching for abstraction: A model. *Mathematical Thinking and Learning*. 12(3), 205-226.
137. Westwood, P. S. (2004). *Learning and learning difficulties: A handbook for teachers*. Camberwell, Victoria: Acer Press
138. Wood, T., & Yackel, E. (1990). The development of collaborative dialogue in small group interactions. In L. P. Steffe & T. Wood (Eds.), *Transforming early childhood mathematics education: An international perspective* (pp. 244-252). Hillsdale, NJ: Erlbaum
139. Yackel, E. & Rasmussen, C. (2002): Beliefs and norms in the mathematics classroom. In: G. Leder, E. Pehkonen & G. Toerner (Eds.), *Beliefs: A hidden variable in mathematics education?* (pp. 313-330). The Netherlands: Kluwer Academic Publishers. MATHDI 2003c.02416
140. Yackel, E.; Cobb, P.; Wood, T.; Wheatley, G. & Merkel, G. (1990): The importance of social interactions in childrens construction of mathematical knowledge. In: T. Cooney (Ed.), *Teaching and learning mathematics in the 1990s*. Reston, VA: National Council of Teachers of Mathematics.
141. Yackel, E.; Rasmussen, C. & King, K. (2001): Social and sociomathematical norms in an advanced undergraduate mathematics course. *Journal of Mathematical Behavior* 19, 1-13. [http://math.coe.uga.edu/olive/EMAT8990FYDS08/Cowner\\_Arg\\_Proof\\_for\\_8990.pdf](http://math.coe.uga.edu/olive/EMAT8990FYDS08/Cowner_Arg_Proof_for_8990.pdf)
142. Yousef, A. (1997) The Effect of the GSP on the Attitude toward Geometry of High School Students. Dissertation Abstract International, A 58105, Ohio University.

## WEBSITES

1. <http://uoi.academia.edu/STAVROULAPATSIOMITOU>
2. [http://www.academia.edu/2054056/Didactic\\_Approaches\\_for\\_teaching\\_Mathematics\\_by\\_using\\_different\\_learning\\_styles](http://www.academia.edu/2054056/Didactic_Approaches_for_teaching_Mathematics_by_using_different_learning_styles)
3. [https://www.researchgate.net/profile/Stavroula\\_Patsiomitou/](https://www.researchgate.net/profile/Stavroula_Patsiomitou/)
4. <http://hcicte2012.uth.gr/main/sites/default/files/proc/Proceedings/Patsiomitou.pdf>
5. [http://www.math.uoa.gr/me/dipl/dipl\\_patsiomitou.pdf](http://www.math.uoa.gr/me/dipl/dipl_patsiomitou.pdf)
6. [http://www.cpre.org/images/stories/cpre\\_pdfs/lp\\_science\\_rr63.pdf](http://www.cpre.org/images/stories/cpre_pdfs/lp_science_rr63.pdf)
7. [http://www.melt.gr/index.php?option=com\\_easygallery&act=categories&cid=23&Itemid=53](http://www.melt.gr/index.php?option=com_easygallery&act=categories&cid=23&Itemid=53)
8. <http://www.moodle.org.nz/>
9. [http://commons.wikimedia.org/wiki/File:Geometric\\_Cremation\\_urn\\_Athens\\_Agora\\_Museum.jpg](http://commons.wikimedia.org/wiki/File:Geometric_Cremation_urn_Athens_Agora_Museum.jpg)
10. [http://www.richardrogers.co.uk/work/buildings/centre\\_pompidou/completed](http://www.richardrogers.co.uk/work/buildings/centre_pompidou/completed)



# A Hybrid Random Forest based Support Vector Machine Classification supplemented by boosting

By Tarun Rao & T.V. Rajinikanth

*Acharya Nagarjuna University, India*

**Abstract** - This paper presents an approach to classify remote sensed data using a hybrid classifier. Random forest, Support Vector machines and boosting methods are used to build the said hybrid classifier. The central idea is to subdivide the input data set into smaller subsets and classify individual subsets. The individual subset classification is done using support vector machines classifier. Boosting is used at each subset to evaluate the learning by using a weight factor for every data item in the data set. The weight factor is updated based on classification accuracy. Later the final outcome for the complete data set is computed by implementing a majority voting mechanism to the individual subset classification outcomes.

**Keywords:** *boosting, classification, data mining, random forest, remote sensed data, support vector machine.*

**GJCST-C Classification:** *G.4*



A HYBRID RANDOM FOREST BASED SUPPORT VECTOR MACHINE CLASSIFICATION SUPPLEMENTED BY BOOSTING

*Strictly as per the compliance and regulations of:*

# A Hybrid Random Forest based Support Vector Machine Classification Supplemented by Boosting

Tarun Rao<sup>α</sup> & T.V. Rajinikanth<sup>σ</sup>

**Abstract-** This paper presents an approach to classify remote sensed data using a hybrid classifier. Random forest, Support Vector machines and boosting methods are used to build the said hybrid classifier. The central idea is to subdivide the input data set into smaller subsets and classify individual subsets. The individual subset classification is done using support vector machines classifier. Boosting is used at each subset to evaluate the learning by using a weight factor for every data item in the data set. The weight factor is updated based on classification accuracy. Later the final outcome for the complete data set is computed by implementing a majority voting mechanism to the individual subset classification outcomes. This method is evaluated and the results compared with that of Support Vector machine classifiers and Random forest classification methods. The proposed hybrid method when applied to classify the plant seed data sets gives better results when compared to traditional random forest and support vector machine classification methods used individually without any compromise in classification accuracy.

**Keywords:** *boosting, classification, data mining, random forest, remote sensed data, support vector machine.*

## 1. INTRODUCTION

Many organizations maintain huge data repositories which store data collected from various sources in different formats. The said data repositories are also known as data warehouses. One of the prominent sources of data is remote sensed data collected via satellites or geographical information systems software's[1].

The data thus collected can be of use in various applications including and not restricted to land use[2] [3], species distribution modeling [4] [5] [6] [7], mineral resource identification[8], traffic analysis[10], network analysis [9] and environmental monitoring systems [11] [12]. Data mining is used to extract information from the said data repositories. The information thus mined can help various stakeholders in an organization in taking strategic decisions. Data can be mined from the data repositories using various methodologies

like anomaly detection, supervised classification, clustering, association rule learning, regression, characterization and summarization and sequential pattern mining. In this paper we shall be applying a hybrid classification technique to classify plant seed remote sensed data.

A lot of research has been undertaken to classify plant functional groups, fish species, bird species etc... [7][13][14].The classification of various species shall help in conserving the ecosystem by facilitating ins predicting of endangered species distribution[15]. It can also help in identifying various resources like minerals, water resources and economically useful trees. Various technologies in this regard have been developed. Machine learning methods, image processing algorithms, geographical information systems tools etc..have added to the development of numerous systems that can contribute to the study of spatial data and can mine relevant information which can be of use in various applications. The systems developed can help constructing classification models that in turn facilitate in weather forecasting, crop yield classification, mineral resource identification, soil composition analysis and also locating water bodies near to the agricultural land.

Classification is the process wherein a class label is assigned to unlabeled data vectors. It can be categorized into supervised and un-supervised classification which is also known as clustering. In supervised classification learning is done with the help of supervisor ie. learning through example. In this method the set of possible class labels is known apriori to the end user. Supervised classification can be subdivided into non-parametric and parametric classification. Parametric classifier method is dependent on the probability distribution of each class. Non parametric classifiers are used when the density function is unknown. Examples of parametric supervised classification methods are Minimal Distance Classifier, Bayesian, Multivariate Gaussian, Support Vector machines, Decision Tree and Classification Tree. Examples of non-parametric supervised classification methods are K- nearest Neighbor, Euclidean Distance, Logistic Regression, Neural Network Kernel Density Estimation, Artificial Neural Network and Multilayer Perceptron. Unsupervised classification is just opposite to the supervised classification i.e. learning is done

**Author α:** Acharya Nagarjuna University, India. His current research interests include Data Mining. e-mail: tarun636@gmail.com

**Author σ:** Sreenidhi Institute of Science and Technology, Hyderabad, India. His current research interests include spatial data mining. Web mining, image Processing and Text mining. e-mail: rajinitv@gmail.com

without supervisor ie. learning from observations. In this method set of possible classes is not known to the end user. After classification one can try to assign a name to that class. Examples of un-supervised classification methods are Adaptive resonance theory(ART) 1, ART 2,ART 3, Iterative Self-Organizing Data Analysis Method, K-Means, Bootstrapping Local, Fuzzy C-Means, and Genetic Algorithm[17]. In this paper we shall discuss about a hybrid classification method. The said hybrid method will make use of support vector machine(SVM) classification, random forest and boosting methods. Later its performance is evaluated against traditional individual random forest classifiers and support vector machines.

A powerful statistical tool used to perform supervised classification is Support Vector machines. Herein the data vectors are represented in a feature space. Later a geometric hyperplane is constructed in the feature space which divides the space comprising of data vectors into two regions such that the data items get classified under two different class labels corresponding to the two different regions. It helps in solving equally two class and multi class classification problem. The aim of the said hyper plane is to maximize its distance from the adjoining data points in the two regions. Moreover, SVM's do not have an additional overhead of feature extraction since it is part of its own architecture. Latest research have proved that SVM classifiers provide better classification results when one uses spatial data sets as compared to other classification algorithms like Bayesian method, neural networks and k-nearest neighbors classification methods[18][19].

In Random forest(RF) classification method many classifiers are generated from smaller subsets of the input data and later their individual results are aggregated based on a voting mechanism to generate the desired output of the input data set. This ensemble learning strategy has recently become very popular. Before RF, Boosting and Bagging were the only two ensemble learning methods used. RF can be applied for supervised classification, unsupervised learning and regression. RF has been extensively applied in various areas including modern drug discovery, network intrusion detection, land cover analysis, credit rating analysis, remote sensing and gene microarrays data analysis etc...[20][21].

Other popular ensemble classification methods are bagging and boosting. Herein the complex data set is divided into smaller feature subsets. An ensemble of classifiers is formed with the classifiers being used to classify data items in each feature subset. The said feature subsets are regrouped together iteratively depending on penalty factor also known as the weight factor applied based on the degree of misclassification in the feature subsets. The class label of data items in the complete data set is computed by aggregating the

individual classification outcomes at each feature subset[22][23].

A hybrid method is being proposed in this paper which makes use of ensemble learning from RF classification and boosting algorithm and SVM classification method. The processed seed plant data is divided randomly into feature subsets. SVM classification method is used to derive the output at each feature subset. Boosting learning method is applied so as to boost the classification adeptness at every feature subset. Later majority voting mechanism is applied to arrive at the final classification result of the original complete data set.

Our next section describes Background Knowledge about Random Forest classifier, SVM and Boosting. In section 3 proposed methodology has been discussed. Performance analysis is discussed in Section 4. Section 5 concludes this work and later acknowledgement is given to the data source followed by references.

## II. BACKGROUND KNOWLEDGE

### a) Overview of SVM Classifier

Support vector machine (SVM) is a statistical tool used in various data mining methodologies like classification and regression analysis. The data can be present either in the form of a multi class or two class problem. In this paper we shall be dealing with a two class problem wherein the seed plant data sets need to be categorized under two class labels one having data sets belonging to North America and the other having data sets belonging to South America. It has been applied in various areas like species distribution, locating mineral prospective areas etc..It has become popular for solving problems in regression and classification, consists of statistical learning theory based heuristic algorithms. The advantage with SVM is that the classification model can be built using minimal number of attributes which is not the case with most other classification methods[24]. In this paper we shall be proposing a hybrid classification methodology to classify seed plant data which would lead to improving the efficiency and accuracy of the traditional classification approach.

The seed plant data sets used in the paper have data sets with known class labels. A classification model is constructed using the data sets which can be authenticated against a test data set and can later be used to predict class labels of unlabeled data sets. Since class labels of data sets are known apriori this approach is categorized as supervised classification. In unsupervised classification method also known as clustering the class label details is not known in advance. Each data vector in the data set used for classification comprises of unique attributes which is used to build the classification model[25][19]. The SVM model can be demonstrated geometrically. As demonstrated in fig 1

SVM is represented by a separating hyper plane  $f(x)$  that geometrically bisects the data space thus dividing it into two diverse regions thus resulting in classification of the input data space into two categories.

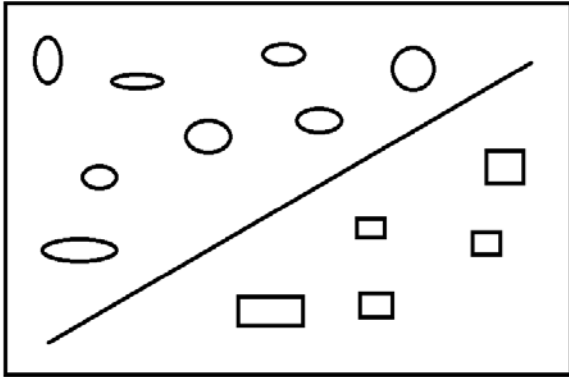


Figure 1 : The Hyperplane

The function  $f(x)$  denotes the hyperplane that separates the two regions and facilitates in classification of the data set. The two regions geometrically created by the hyperplane correspond to the two categories of data under two class labels. A data point  $x_n$  belongs to either of the region depending on the value of  $f(x_n)$ . If  $f(x_n) > 0$  it belongs to one region and if  $f(x_n) < 0$  it belongs to another region. There are many such hyperplanes which can split the data into two regions. But SVM ensures that it selects the hyperplane that is at a maximum distance from the nearest data points in the two regions. There are only few hyperplanes that shall satisfy this criterion. By ensuring this condition SVM provides accurate classification results[27].

SVM's can be represented mathematically as well. Assume that the input data consists of  $n$  data vectors where each data vector is represented by  $x_i \in R_n$ , where  $i (=1, 2, \dots, n)$ . Let the class label that needs to be assigned to the data vectors to implement supervised classification be denoted by  $y_i$ , which is  $+1$  for one category of data vectors and  $-1$  for the other category of data vectors. The data set can be geometrically separated by a hyperplane. Since the hyperplane is represented by a line it can also be mathematically represented by[8][3][28]:

$$\begin{aligned} mx_i + b &\geq +1 \\ mx_i + b &\leq -1 \end{aligned} \tag{1}$$

The hyperplane can also be represented mathematically by [31][32][33]:

$$\begin{aligned} f(x) &= \text{sgn}(mx + b) \\ &= \text{sgn}\left(\sum_{i=1}^n \alpha_i y_i x_i \cdot x + b\right) \end{aligned} \tag{2}$$

where  $\text{sgn}()$  is known as a sign function, which is mathematically represented by the following equation:

$$\text{sgn}(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases} \tag{3}$$

The data vectors are said to be optimally divided by the hyperplane if the distance amid the adjoining data vectors in the two different regions from the given hyperplane is maximum.

This concept can be illustrated geometrically as in Figure 2, where the distance between the adjoining data points close to the hyperplane and the hyperplane is displayed[29][30][28].

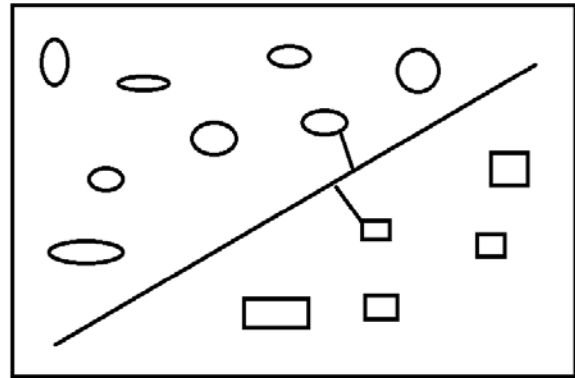


Figure 2 : Distance of the nearest data vectors from the Hyperplane

The distance  $d$  of a data point  $x$  from the hyperplane is represented mathematically by the equation:

$$d = \frac{|(m \cdot x) + b|}{|m|} \tag{4}$$

This hyperplane which has maximum distance  $d$  from adjoining points is computed to implement the said classification. This SVM can be represented as a primal formulation given by the equation [8][5][31]:

$$\begin{aligned} h(m) &= \frac{1}{2} \|m\|^2 + \text{Training error} \\ \text{subject to } y_i(m \cdot x_i + b) &> 1, \forall i \end{aligned} \tag{5}$$

The idea is to increase the margin and reduce the training error. The data sample records in the training data set belong to input set. Each of the data vectors have precise attributes based on which the classification model is built. These set of attributes are said to form a feature space. The kernel function bridges the gap between the feature space and the input space and enables to carry out classification on input space rather than complicated feature space. [29].

In this paper we have used Gaussian radial basis functions (RBF). SVM's make use of the radial basis kernel function to be able to work at the simpler input space level. The RBF kernel used is represented mathematically by[3][29]:

$$K(x_1, x_2) = \exp\left(-\frac{|x_1 - x_2|^2}{2\sigma^2}\right) \tag{6}$$

SVM selects a local Gaussian function and later the global Gaussian function is computed by aggregating all the local Gaussian function.

SVM can be used to solve either two class or multi-class problems. Multiclass classification problems



can be solved using various methods. One method is to move the data vectors to a different space thereby making the problem linear. The other method is to split the multi class problem into numerous two class problems and later with a voting mechanism combine the solutions of individual two class problems to get the solution of the original multi class problem. [8].

The steps followed while using SVM in classifying data are mentioned in the below algorithm [16]:

-----  
Algorithm 2 Classification using linear kernel based SVM Classifier  
-----

Input: I: Input data  
Output: V: Support vectors set  
begin  
Step 1: Divide the given data set into two set of data items having different class labels assigned to them  
Step 2: Add them to support vector set V  
Step 3: Loop the divided n data items  
Step 4: If a data item is not assigned any of the class labels then add it to set V  
Step 5: Break if insufficient data items are found  
Step 6: end loop  
Step7: Train using the derived SVM classifier model and test so as to validate over the unlabelled data items.  
end  
-----

#### b) Overview of Random Forest Classifier

Ensemble learning algorithms use an ensemble or a group of classifiers to classify data. Hence they give more accurate results as compared to individual classifiers. Random forest classifier is an example for ensemble classifiers. Random forests make use of an ensemble of classification trees [34][35][36][37][38][41].

In RF classification method the input data set is first subdivided into two subsets, one containing two thirds of the data points and the other containing the remaining one third. Classification tree models are constructed using the subset comprising of two thirds of data points The subset which contains one third data of data points which are not used at any given point of time to construct classification trees and are used for validation are called out of bag(OOB) data samples of the trees. There is no truncation applied at every classification tree. Hence every classification tree used in RF classification method is maximal in nature. Later RF classification method follows a majority voting process wherein classification output of every classification tree casts a vote to decide the final outcome of the ensemble classifier ie.. assigning a class label to a data item x[21]. The set of features are used to create a classification tree model at every randomly chosen subset[37]. This set of features shall remain constant throughout the growing of random forest.

In RF, the test set is used to authenticate the classification results and also used for predicting the class labels for unlabeled data after the classification model is built. It also helps in cross validation of results among different classification results provided by various classification trees in the ensemble. To perform the said cross validation the out of bag(OOB) samples are used.. The individual classification tree outcomes are aggregated with a majority vote and the cumulative result of the whole ensemble shall be more accurate and prone to lesser classification error than individual classification tree results[26].

Every classification tree in the random forest ensemble is formed using the randomly selected two thirds of input variables, hence there is little connection between different trees in the forest. One can also restrict the number of variables that split a parent node in a classification tree resulting in the reduction of connection between classification trees. The Random forest classification method works better even for larger data sets. This is not the case with other ensemble methods[1][2]. In this paper we shall be using the both boosting and random forest ensemble classification methods along with support vector machines to give a more accurate classification output. This hybrid method shall be more robust to noise as compared to individual classification method.

RF classification method works with both discreet and continuous variables which is not the case with other statistical classification modeling methods. Furthermore, there is no limit on the total number of classification trees that are generated in the ensemble process and the total number of variable or data samples(generally two thirds are used) in every random subset used to build the classification trees[36].

RF rates variables based on the classification accuracy of the said variable relative to other variables in the data set. This rank is also known as importance index. It reflects the relative importance of every variable in the process of classification. The importance index of a variable is calculated by averaging the importance of the variable across classification trees generated in the ensemble. The more the value of this importance index, the greater is a variables importance for classification. Another parameter obtained by dividing the variable's importance index by standard error is called z-score. Both importance index as well as z-score play a significant role in ensuring the efficiency of the classification process[25][36][39][38].

The importance of a variable can also be assessed by using two parameters, Gini Index decrease and OOB error estimation. Herein relative importance of variables are calculated which is beneficial in studies wherein the numbers of attributes are very high and thus leading to relative importance gaining prominence[40]. For any training data set T Gini index value can be computed for two variables i and j as[2]:

$$\sum \sum_{j \neq i} \left( \frac{k(C_i, X)}{|X|} \right) \cdot \left( \frac{k(C_j, X)}{|X|} \right) \quad (7)$$

where  $\frac{k(C_i, X)}{|X|}$  is the probability that a selected case belongs to class  $C_i$ .

RF method provides precise results with respect to variation and bias[39]. The performance of the RF classification method is better compared to other classifiers like support vector machines, Neural Networks and discriminant analysis. In this paper a hybrid classification method coalescing the advantages of both Random forest and Support vector machines in addition to boosting is used. The RF algorithm is becoming gradually popular with applications like forest classification, credit rate analysis, remote sensing image analysis, intrusion detection etc.

Yet another parameter that can contribute in assessing the classification is proximity measure of two samples. The proximity measure is the number of classification trees in which two data samples end up in the same node. This parameter when divided by the number of classification trees generated can facilitate in detecting outliers in the data sets. This computation requires large amount of memory space, depending on the total number of sample records and classification trees in the ensemble[1]. The pseudo code for Random Forest algorithm is mentioned below[42]:

-----  
 Random Forest Algorithm:  
 -----

- Input: D: training sample  
 a: number of input instance to be used to generate classification tree  
 T: total number of classification trees in random forest  
 OT: Classification Output from each tree T  
 1) OT is empty  
 2) for  $i=1$  to T  
 3)  $D_b$  = Form random sample subsets after selecting 2/3rd instances randomly from D  
 /\* For every tree this sample would be randomly selected\*/  
 4)  $C_b$  = Build classification trees using random subsets  $D_b$   
 5) Validate the classifier  $C_b$  using remaining 1/3rd instances //Refer Step 3.  
 6) OT=store classification outputs of classification trees  
 7) next i  
 8) Apply voting mechanism to derive output ORT of the Random forest(ensemble of classification trees)  
 9) return ORT  
 -----

c) *Overview of Boosting*

Ensemble learning is a process wherein a data set is divided into subsets. Individual learners are then used to classify and build the model for each of these subsets. Later the individual learning models are combined so as to determine the final classification

model of the complete data set. As the complex large data set is divided into smaller random subsets and classification model is applied on these smaller subsets the said process of ensemble learning results in improving classification efficiency and gives more accurate results. Numerous classification methodologies like bagging, boosting etc...can also be used in learning by constructing an ensemble[43][44][45].

In this research paper boosting method has been used to create the said ensemble. It works by rewarding successful classifiers and by applying penalties to unsuccessful classifiers. In the past it has been used in various applications like machine translation [46], intrusion detection [47], forest tree regression, natural language processing, unknown word recognition [48] etc.

Boosting is applied to varied types of classification problems. It is an iterative process wherein the training data set is regrouped together into subsets and various classifiers are used to classify data samples in the subsets. The data samples which were difficult to classify by a classifier also known as a weak learner at one stage are classified using new classifiers that get added to the ensemble at a later stage[49][50][51]. In this way at each stage a new classifier gets augmented to the ensemble. The difficulty in classifying a data item  $X_i$  at stage  $k$  is represented by a weight factor  $W_k(i)$ . The regrouping of training sets at each step of learning is done depending on the weight factor  $W_k(i)$ [22]. The value of the weight factor is proportional to the misclassification of the data. This way of forming regrouped data samples at every stage depending on the weight factor is called re-sampling version of boosting. Yet another way of implementing boosting is by reweighting wherein weight factor is assigned iteratively to every data item in the data set and the complete data set is used at every subsequent iteration by modifying the weights at every stage[48][52].

The most popular boosting algorithm called Adaboost[23]. Adaboost stands for Adaptive Boosting. It adapts or updates weights of the data items based on misclassification of training samples due to weak learners and regroups the data subsets depending on the new weights. The steps of Adaboost algorithm is mentioned below:

-----  
 Adaboost Algorithm  
 -----

- Divide the data set into random subsets  
 Set uniform example weights to all data items in the subsets  
 For each base learner  
 do  
 Train base learner with weighted sample of a sample data set  
 Test base learner on complete data set.  
 Update learner weight based on



misclassifications by base learners also called weak learners if misclassification occurs  
 Set example weights based on predictions by ensemble of classifiers.  
 end for

In the next section the proposed hybrid methodology is discussed in detail.

### III. PROPOSED METHODOLOGY

In this paper we shall construct a hybrid classification model which shall facilitate in predicting the class label of seed plant data from test data sets. The methodology recommended has been denoted as a schematic diagram as mentioned in Fig 3 and the detailed explanation of the steps followed has been given in the following subsections.

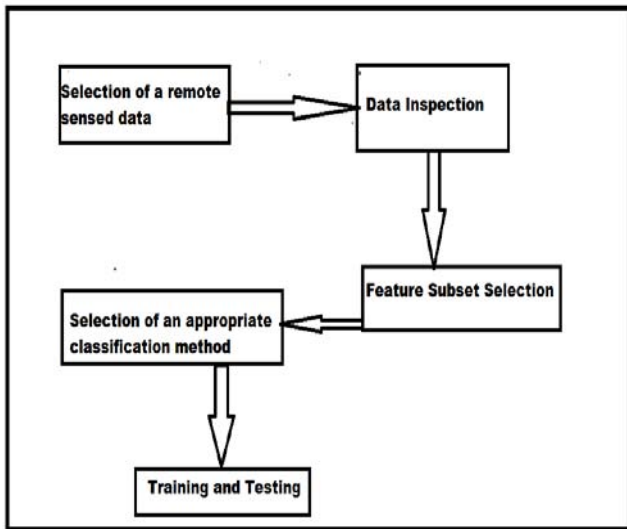


Figure 3 : Proposed Model

a) *Selection of Remote sensed data*

The data sets collected in this paper belong to various types of seed plant family viz Pinopsida, Dicotyledoneae, Monocotyledoneae, Cycadopsida, Pinopsida, Gnetopsida, Lycopodiopsida, Agaricomycetes and Marchantiopsida.

b) *Data Inspection*

The seed plant data sets are pre-processed and any missing values or duplicate values are eliminated by either ignoring tuples comprising of duplicate values or by manually entering values or replacing with a global constant or a mean value into tuples with missing values[53].

c) *Feature subset Selection*

The data sets are randomly divided into n different random subsets each subset comprising of two third of the whole data set. Classification methods are applied to each of these random subsets. The remaining

one third data sets at each subsets is used as a test set. At each random subset the following attributes were used so as to implement the classification method discussed in the next subsection: id, continent, specificEpithet and churn. Now churn is a variable that is set to yes if the seed plant data belongs to North America or if it belongs to South America it is set to no.

d) *Selection of an appropriate classification method*

In this paper seed plant data sets are classified using a hybrid classification method which makes use of Random forest, SVM classifier and boosting ensemble learning method. In the hybrid methodology the input data set is randomly subdivided into subsets. Each data item in each of the subset has a weight factor associated with it. The data items in the subsets are classified by SVM classifier. If a misclassification has occurred then the weight factor of the data items is increased otherwise it is reduced. The data subsets are rearranged and again SVM classifier is used to perform classification at each subset. The weights are again updated depending on whether it is a proper classification or a misclassification. These steps are iteratively repeated till all the weights get updated to a very low value. The output of the input data set is computed by applying voting mechanism to all the random subsets classification outputs[34]. The algorithm for the proposed hybrid methodology is given in the sample code herein:

Algorithm 1 Hybrid classification using RF and SVM supplemented by boosting

```

    Input: D Training Instances
    Intermediate Output: Osvm, Classification output at each feature subset
    Output: O, Classification Output for the hybrid method
    Step 1)Begin
    Step 2)Initialize the weight  $w_i$  for each data vector  $i \in D$ .
    Step 3) Generate a new data feature subset  $D_i$  from D using random replacement method.
    Step 4)begin
    Step 5)For each random feature subset  $D_i$  do
    Step 6)begin
    Step 7)Apply SVM to each feature subset
    Step 8) Generate  $O_{svm}$ , the classification output from
    Step 9) end
    Step 10) Update the weights of all the data vectors in the training set depending on the classification outcome. If an example was misclassified then its weight is increased, or else the weight is decreased.
    Step 11) Repeat steps 2 to 10 by regenerating random subsets till all the input data vectors are appropriately classified or apply iteration limit.
    Step 12) Compute output O of the complete data set by applying majority voting mechanism among the final outputs of each of the Random feature subsets  $D_i$  of The original set D obtained after Step 11.
    Step 13) Return O
    Step 14)End
    
```

e) *Training and Testing*

The obtained classification output at each random subset is validated by using the hybrid classifier model to test against the complete data set.

In this paper 10 random feature subsets were used and at every subset SVM classifier was used to perform the said classification. Voting mechanism was then applied to derive the final classification output. In this paper a total of 180 support vectors were used.

IV. PERFORMANCE ANALYSIS

a) *Environment Setting*

The study area included is from North and South America. It includes data pertaining to localities wherein seed plant species are present.

A total of 599 data set records from North American region and a total of 401 data set records from South American region are analyzed in order to execute the proposed method. Sample records used in this paper are shown in Table I shown below:

Table 1 : Sample records

id	higherGeography	continent	family	scientificName	decimalLatitude	decimalLongitude	genus	specificEpi	chum
2759	North America,	North	Lycoperdaceae	Calvatiaarctica	72	-40	Calvati	arctica	yes
86	GREENLAND	America	aeae	Empetrum			aeae		
3333	North America,	North	Ericaceae	&Wiegand	52	-56	Empetrum	eamesii	yes
01	North America,	America	Ranunculaceae	Thalictrum			Thalictrum	terrae-novae	
2717	North America,	North	Greene	Thalictrum	52	-56	Thalictrum	terrae-novae	yes
58	North America,	America							

A total of 3 features are first extracted as stated in section III. Then obtained feature vectors are fed into the hybrid classifier whose results are compared against SVM and Random forest Classifier results. A total of 80 data set records act as test data set and are used to authenticate the classification results obtained. The proposed method has been implemented under the environment setting as shown in Table II[54].

Table 2 : Environment Setting

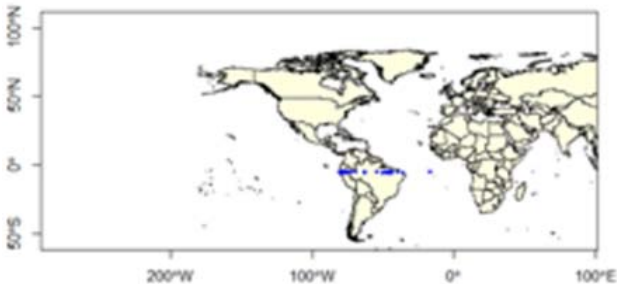
Item	Capacity
CPU	Intel CPU G645 @2.9 GHz processor
Memory	8GB RAM
OS	Windows 7 64-bit
Tools	R, R Studio

b) *Result Analysis*

Classification of the spatial data sets can be represented as a confusion or error matrix view as shown in Table III. And the classified seed plant data is demonstrated in Figure 4[54].

Table 3 : Confusion / Error Matrix View

Real group	Classification result	
	North America	South America
North America	True Negative(TN)	False Positive(FP)
South America	False Negative(FN)	True Positive(TP)



(a)

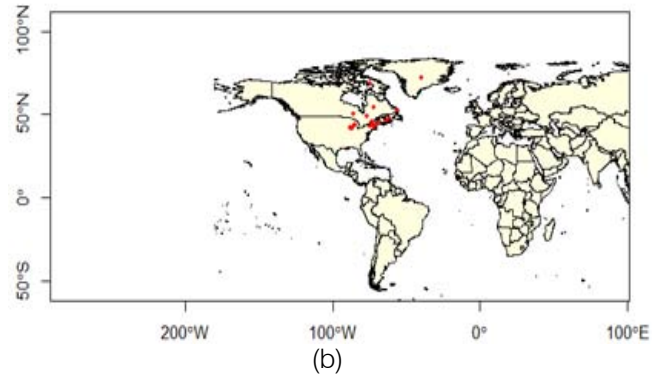


Figure 4 : Plot of seed plant data (a) Seed plant belonging to South America marked in blue. (b) Seed plant belonging to North America marked in red.

It is observed that the most conventionally utilized evaluation metrics in classification are accuracy, specificity, positive predictive value and negative predictive value. The formulae for accuracy, specificity, prevalence and negative predictive value are provided by equations (8), (9), (10) and (11)[54]:

$$\text{Accuracy} = \frac{TP + TN}{(TP + FN + FP + TN)} \times 100 \quad (8)$$

$$\text{Specificity} = \frac{TN}{(TN + FP)} \times 100 \quad (9)$$

$$\text{Prevalence} = \frac{(TN + FN)}{(TP + FP + TN + FN)} \times 100 \quad (10)$$

$$\text{Neg. Predictive Value} = \frac{TN}{TN + FN} \times 100 \quad (11)$$

The efficiency of the proposed hybrid classification is evaluated and compared with traditional RF ensemble and SVM classification methods. The confusion or error matrix view for hybrid classifier is given in Table IV.

Table 4 : Confusion Matrix for Hybrid Classifier

Prediction	Reference	
	South America	North America
South America	1	4
North America	5	70

The confusion matrix or error matrix view for SVM Classifier is given in Table V and for RF Classifier in Table VI.

Table 5 : Confusion Matrix for SVM

Prediction	Reference	
	South America	North America
South America	8	7
North America	49	16

Table 6 : Confusion Matrix for RF

Prediction	Reference	
	South America	North America
South America	36	11
North America	21	12

Performance Measures using evaluation metrics are specified in Fig 5 which are calculated using equations (8), (9), (10)and (11).

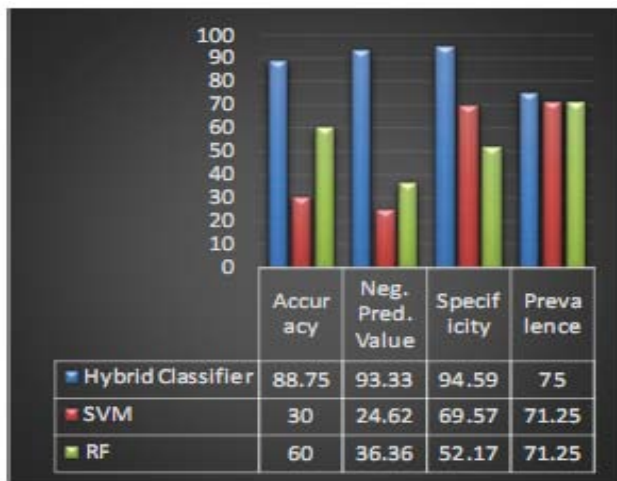


Figure 5 : Performance measures for proposed hybrid, SVM,RF classifiers

Hence, from the results obtained it is inferred that the proposed hybrid classifier is better than traditional SVM and random forest classifiers.

### V. CONCLUSION

In this paper hybrid classifier based on random forest, SVM and boosting methods is used to classify seed plant data. The hybrid classification results are compared with the results attained by implementing classification using traditional SVM and RF classifiers. The research has established that the hybrid approach of classification is more efficient as compared to traditional SVM and RF classifiers since it gives higher values of

accuracy, specificity, positive predictive value and negative predictive value.

The reason for better results in the case of hybrid classification methodology used in this paper is since it makes use of the advantages of each of the individual traditional SVM, RF classifications methods. Furthermore, the classification results are supplemented using boosting ensemble classification method. In the future the proposed method can be used so as to classify vector, raster remote sensed data that can be collected via satellites and various geographical information systems.

### VI. ACKNOWLEDGMENT

We direct our frank appreciativeness to the Field Museum of Natural History(Botany)- Seed Plant Collection (accessed through GBIF data portal, <http://data.gbif.org/datasets/resource/14346,2013-06-03>) for providing us with different seed plant data sets. We also thank ANU university for providing all the support in the work conducted.

### REFERENCES RÉFÉRENCES REFERENCIAS

1. Pall Oskar Gislason, Jon Atli Benediktsson, Johannes R. Sveinsson, Random Forests for land cover classification, Pattern Recognition Letters, Volume 27, Issue 4, March 2006, Pages 294-300, ISSN 0167-8655,<http://dx.doi.org/10.1016/j.patrec.2005.08.011>.
2. V.F. Rodriguez-Galiano, B.Ghimire, J. Rogan, M. Chica-Olmo, J.P .Rigol-Sanchez, An assessment of the effectiveness of a random forest classifier for land-cover classification, ISPRS Journal of Photogrammetry and Remote Sensing, Volume 67, January 2012, Pages 93-104, ISSN0924-2716, <http://dx.doi.org/10.1016/j.isprsjprs.2011.11.002>.
3. Hassiba Nemmour, Youcef Chibani, Multiple support vector machines for land cover change detection: An application for mapping urban extensions, ISPRS Journal of Photogrammetry and Remote Sensing, Volume 61, Issue 2, November 2006, Pages 125-133,ISSN0924-2716,<http://dx.doi.org/10.1016/j.isprsjprs.2006.09.004>.
4. P.A. Aguilera, A. Fernández, F.Reche, R. Rumí, Hybrid Bayesian network classifiers: Application to species distribution models, Environmental Modelling & Software, Volume25, Issue12, December 2010, Pages 1630-1639,ISSN1364 8152,<http://dx.doi.org/10.1016/j.envsoft.2010.04.016>.
5. Till Rumpf, Christoph Römer, Martin Weis, Markus Sökefeld, Roland Gerhards, Lutz Plümer, Sequential support vector machine classification for small-grain weed species discrimination with special regard to Cirsiumarvense and Galiumaparine, Computers and Electronics in Agriculture, Volume 80, January 2012,

- Pages89-96, ISSN0168-1699,<http://dx.doi.org/10.1016/j.compag.2011.10.018>.
6. Jing Hu, Daoliang Li, Qingling Duan, Yueqi Han, Guifen Chen, XiuliSi, Fish species classification by color, texture and multi-class support vector machine using computer vision, *Computers and Electronics in Agriculture*, Volume 88, October 2012, Pages 133-140, ISSN0168-1699, <http://dx.doi.org/10.1016/j.compag.2012.07.008>.
  7. L. Naidoo, M.A. Cho, R. Mathieu, G. Asner, Classification of savanna tree species, in the Greater Kruger National Park region, by integrating hyperspectral and LiDAR data in a Random Forest data mining environment, *ISPRS Journal of Photogrammetry and Remote Sensing*, Volume 69, April 2012, Pages 167-179, ISSN0924-2716, <http://dx.doi.org/10.1016/j.isprsjprs.2012.03.005>.
  8. Maysam Abedi, Gholam-Hossain Norouzi, Abbas Bahroudi, Support vector machine for multi-classification of mineral prospectivity areas, *Computers & Geosciences*, Volume 46, September2012, Pages 272-283, ISSN 0098-004, <http://dx.doi.org/10.1016/j.cageo.2011.12.014>.
  9. D.Ruano- Ordás, J. Fdez-Glez, F. Fdez-Riverola, J.R. Méndez, Effective scheduling strategies for boosting performance on rule-based spam filtering frameworks, *Journal of Systems and Software*, Volume86, Issue12, December2013, Pages 3151-3161, ISSN 0164-1212, <http://dx.doi.org/10.1016/j.jss.2013.07.036>.
  10. N. Abdul Rahim, Paulraj M.P., A.H .Adom, Adaptive Boosting with SVM Classifier for Moving Vehicle Classification, *Procedia Engineering*, Volume 53, 2013, Pages 411-419, ISSN1877-7058, <http://dx.doi.org/10.1016/j.proeng.2013.02.054>.
  11. Konstantinos Topouzelis, Apostolos Pysillos, Oil spill feature selection and classification using decision tree forest on SAR image data, *ISPRS Journal of Photogrammetry and Remote Sensing*, Volume 68, March 2012, Pages 135-143, ISSN 0924-2716, <http://dx.doi.org/10.1016/j.isprsjprs.2012.01.005>.
  12. Yang Shao, Ross S. Lunetta, Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points, *ISPRS Journal of Photogrammetry and Remote Sensing*, Volume70, June2012, Pages78-87, ISSN0924-2716, <http://dx.doi.org/10.1016/j.isprsjprs.2012.04.001>.
  13. Paul Bosch, Julio López, Héctor Ramírez, Hugo Robotham, Support vector machine under uncertainty: An application for hydroacoustic classification of fish-schools in Chile, *Expert Systems with Applications*, Volume 40, Issue 10, August 2013 ,Pages 4029-4034, ISSN0957-4174, <http://dx.doi.org/10.1016/j.eswa.2013.01.006>.
  14. Hongji Lin, Han Lin, Weibin Chen, Study on Recognition of Bird Species in Minjiang River Estuary Wetland, *Procedia Environmental Sciences*, Volume10, Part C, 2011, Pages 2478-2483, ISSN1878-0296, <http://dx.doi.org/10.1016/j.proenv.2011.09.386>.
  15. Rafael Pino-Mejías, María Dolores Cubiles-de-la-Vega, María Anaya-Romero, Antonio Pascual-Acosta, Antonio Jordán-López, Nicolás Bellinfante-Crocci, Predicting the potential habitat of oaks with data mining models and the R system, *Environmental Modelling & Software*, Volume 25 , Issue 7, July 2010, Pages 826-836, ISSN 1364-8152, <http://dx.doi.org/10.1016/j.envsoft.2010.01.004>.
  16. S.N.Jeyanthi, Efficient Classification Algorithms using SVMs for Large Datasets, A Project Report Submitted in partial fulfillment of the requirements for the Degree of Master of Technology in Computational Science, Supercomputer Education and Research Center, IISC, BANGALORE, INDIA, June 2007.
  17. Yugal kumar, G. Sahoo, Analysis of Parametric & Non Parametric Classifiers for Classification Technique using WEKA, *I.J. Information Technology and Computer Science*, 2012, 7, 43-49 Published Online July 2012 in MECS DOI:10.5815/ijitcs.2012.07.06
  18. M.Arun Kumar, M. Gopal, A hybrid SVM based decision tree, *Pattern Recognition*, Volume 43, Issue 12, December 2010, Pages 3977-3987, ISSN0031-3203, <http://dx.doi.org/10.1016/j.patcog.2010.06.010>.
  19. Rajasekhar. N.; Babu, S.J.; Rajinikanth, T.V., "Magnetic resonance brain images classification using linear kernel based Support Vector Machine," *Engineering (NUICONe)*, 2012 Nirma University International Conference on , vol., no pp.1,5,68 Dec.2012 doi10.1109/NUICONe.2012.6493213
  20. V.F. Rodríguez-Galiano, F.Abarca-Hernández, B. Ghimire, M. Chica-Olmo, P.M .Akinson, C. Jeganathan, Incorporating Spatial Variability Measures in Land-cover Classification using Random Forest, *Procedia Environmental Sciences*, Volume3, 2011, Pages44-49, ISSN1878-0296, <http://dx.doi.org/10.1016/j.proenv.2011.02.009>.
  21. Reda M. Elbasiony, Elsayed A.Sallam, Tarek E. Eltobely, Mahmoud M. Fahmy, A hybrid network intrusion detection framework based on random forests and weighted k-means, *A in Shams Engineering Journal*, Available online 7 March 2013, ISSN 2090-4479, <http://dx.doi.org/10.1016/j.asej.2013.01003>.
  22. Gonzalo Martínez-Muñoz, Alberto Suárez, Using boosting to prune bagging ensembles, *Pattern Recognition Letters*, Volume28, Issue 1, 1 January 2007, Pages 156-165, ISSN0167-8655, <http://dx.doi.org/10.1016/j.patrec.2006.06.018>.
  23. Chun-Xia Zhang, Jiang-She Zhang, Gai-Ying Zhang, An efficient modified boosting method for solving classification problems, *Journal of Computational*

- and Applied Mathematics, Volume 214, Issue 2, 1 May 2008, Pages 381-392, ISSN 0377-0427, <http://dx.doi.org/10.1016/j.cam.2007.03.003>.
24. F. Löw, U. Michel, S. Dech, C. Conrad, Impact of feature selection on the accuracy and spatial uncertainty of per-field crop classification using Support Vector Machines, *ISPRS Journal of Photogrammetry and Remote Sensing*, Volume 85, November 2013, Pages 102-119, ISSN 0924-2716, <http://dx.doi.org/10.1016/j.isprsjprs.2013.08.007>.
  25. Liu Mingjun Wang, Jun Wang, Duo Li, Comparison of random forest, support vector machine and back propagation neural network for electronic tongue data classification: Application to the recognition of orange beverage and Chinese vinegar, *Sensors and Actuators B: Chemical*, Volume 177, February 2013, Pages 970-980, ISSN 0925-4005, <http://dx.doi.org/10.1016/j.snb.2012.11.071>.
  26. Ching-Chiang Yeh, Der-Jang Chi, Yi-Rong Lin, Going-concern prediction using hybrid random forests and rough set approach, *Information Sciences*, Available online 6 August 2013, ISSN 0020-0255, <http://dx.doi.org/10.1016/j.ins.2013.07.011>.
  27. Hsun-Jung Cho, Ming-Te Tseng, A support vector machine approach to CMOS-based radar signal processing for vehicle classification and speed estimation, *Mathematical and Computer Modelling*, Volume 58, Issues 1-2, July 2013, Pages 438-448, ISSN 0895-7177, <http://dx.doi.org/10.1016/j.mcm.2012.11.003>.
  28. Lam Hong Lee, Rajprasad Rajkumar, Lai Hung Lo, Chin Heng Wan, Dino Isa, Oil and gas pipeline failure prediction system using long range ultrasonic transducers and Euclidean-Support Vector Machines classification approach, *Expert Systems with Applications*, Volume 40, Issue 6, May 2013, Pages 1925-1934, ISSN 0957-4174, <http://dx.doi.org/10.1016/j.eswa.2012.10.006>.
  29. David Meyer, Support Vector Machines The Interface to lib svm in package e1071, Technische University of Wien, Austria, September, 2012.
  30. Xiaowei Yang, Qiaozhen Yu, Lifang He, Tengjiao Guo, The one-against-all partition based binary tree support vector machine algorithms for multi-class classification, *Neurocomputing*, Volume 113, 3 August 2013, Pages 1-7, ISSN 0925-2312, [10.1016/j.neucom.2012.12.048](http://dx.doi.org/10.1016/j.neucom.2012.12.048).
  31. Steve R. Gunn, Support Vector Machines for Classification and Regression, A Technical Report Faculty of Engineering, Science and Mathematics School of Electronics and Computer Science, University Of South Hampton, May 1998.
  32. Asdrúbal López Chau, Xiaou Li, Wen Yu, Convex and concave hulls for classification with support vector machine, *Neuro computing*, Volume 122, 25 December 2013, Pages 198-209, ISSN 0925-2312, <http://dx.doi.org/10.1016/j.neucom.2013.05.040>.
  33. Xinjun Peng, Yifei Wang, Dong Xu, Structural twin parametric-margin support vector machine for binary classification, *Knowledge-Based Systems*, Volume 49, September 2013, Pages 63-72, ISSN 0950-7051, <http://dx.doi.org/10.1016/j.knsys.2013.04.013>.
  34. Adnan Idris, Muhammad Rizwan, Asifullah Khan, Churn prediction in telecom using Random Forest and PSO based data balancing in combination with various feature selection strategies, *Computers & Electrical Engineering*, Volume 38, Issue 6, November 2012, Pages 1808-1819, ISSN 0045-7906, <http://dx.doi.org/10.1016/j.compeleceng.2012.09.001>.
  35. Elias Zintzaras, Axel Kowald, Forest classification trees and forest support vector machines algorithms: Demonstration using microarray data, *Computers in Biology and Medicine*, Volume 40, Issue 5, May 2010, Pages 519-524, ISSN 0010-4825, <http://dx.doi.org/10.1016/j.combiomed.2010.03.006>.
  36. Ching-hiang Yeh, Fengyi Lin, Chih-Yu Hsu, A hybrid KMV model, random forests and rough set theory approach for credit rating, *Knowledge-Based Systems*, Volume 33, September 2012, Pages 166-172, ISSN 0950-7051, <http://dx.doi.org/10.1016/j.knsys.2012.04.004>.
  37. Björn Waske, Sebastian vander Linden, Carsten Oldenburg, Benjamin Jakimow, Andreas Rabe, Patrick Hostert, imageRF – A user-oriented implementation for remote sensing image analysis with Random Forests, *Environmental Modeling & Software*, Volume 35, July 2012, Pages 192-193, ISSN 1364-8152, <http://dx.doi.org/10.1016/j.envsoft.2012.01.014>.
  38. Miao Liu, Mingjun Wang, Jun Wang, Duo Li, Comparison of random forest, support vector machine and back propagation neural network for electronic tongue data classification: Application to the recognition of orange beverage and Chinese vinegar, *Sensors and Actuators B: Chemical*, Volume 177, February 2013, Pages 970-980, ISSN 0925-4005, <http://dx.doi.org/10.1016/j.snb.2012.11.071>.
  39. Robin Genuer, Jean-Michel Poggi, Christine Tuleau-Malot, Variable selection using r random forests, *Pattern Recognition Letters*, Volume 31, Issue 14, 15 October 2010, Pages 2225-2236, ISSN 0167-8655, <http://dx.doi.org/10.1016/j.patrec.2010.03.014>.
  40. Katherine R. Gray, Paul Aljabar, Rolf A. Heckemann, Alexander Hammers, Daniel Rueckert, for the Alzheimer's Disease Neuroimaging Initiative, Random forest-based similarity measures for multimodal classification of Alzheimer's disease, *NeuroImage*, Volume 65, 15 January 2013, Pages 167-175, ISSN 10538119, <http://dx.doi.org/10.1016/j.neuroimage.2012.09.065>.
  41. Meng Xiao, Hong Yan, Jinzhong Song, Yuzhou Yang, Xianglin Yang, Sleep stages classification

- based on heart rate variability and random forest, *Biomedical Signal Processing and Control*, Volume 8, Issue 6, November 2013, Pages 624-633, ISSN 1746-8094, <http://dx.doi.org/10.1016/j.bspc.2013.06.001>.
42. Akin Özçift, Random forests ensemble classifier trained with data resampling strategy to improve cardiac arrhythmia diagnosis, *Computers in Biology and Medicine*, Volume 41, Issue 5, May 2011, Pages 265-271, ISSN 0010-4825, <http://dx.doi.org/10.1016/j.complbiomed.2011.03.001>.
  43. Simone Borra, Agostino Di Ciaccio, Improving nonparametric regression methods by bagging and boosting, *Computational Statistics & Data Analysis*, Volume 38, Issue 4, 28 February 2002, Pages 407-420, ISSN 0167-9473, [http://dx.doi.org/10.1016/S0167-9473\(01\)00068-8](http://dx.doi.org/10.1016/S0167-9473(01)00068-8).
  44. Imed Zitouni, Hong-Kwang Jeff Kuo, Chin-Hui Lee, Boosting and combination of classifiers for natural language call routing systems, *Speech Communication*, Volume 41, Issue 4, November 2003, Pages 647-661, ISSN 0167-6393, [http://dx.doi.org/10.1016/0167-6393\(03\)00103-1](http://dx.doi.org/10.1016/0167-6393(03)00103-1).
  45. Jafar Tanha, Maarten van Someren, Hamideh Afsarmanesh, Boosting for Multiclass Semi-Supervised Learning, *Pattern Recognition Letters*, Available online 21 October 2013, ISSN 0167-8655, <http://dx.doi.org/10.1016/j.patrec.2013.10.008>.
  46. Tong Xiao, Jingbo Zhu, Tongran Liu, Bagging and Boosting statistical machine translation systems, *Artificial Intelligence*, Volume 195, February 2013, Pages 496-527, ISSN 0004-3702, <http://dx.doi.org/10.1016/j.artint.2012.11.005>.
  47. Tansel Özyer, Reda Alhajj, Ken Barker, Intrusion detection by integrating boosting genetic fuzzy classifier and data mining criteria for rule pre-screening, *Journal of Network and Computer Applications*, Volume 30, Issue 1, January 2007, Pages 99-113, ISSN 1084-8045, <http://dx.doi.org/10.1016/j.jnca.2005.06.002>.
  48. Jakkrit TeCho, Cholwich Nattee, Thanaruk Theeramunkong, Boosting-based ensemble learning with penalty profiles for automatic Thai unknown word recognition, *Computers & Mathematics with Applications*, Volume 63, Issue 6, March 2012, Pages 1117-1134, ISSN 0898-1221, <http://dx.doi.org/10.1016/j.camwa.2011.11.062>.
  49. Chun-Xia Zhang, Jiang-She Zhang, A local boosting algorithm for solving classification problems, *Computational Statistics & Data Analysis*, Volume 52, Issue 4, 10 January 2008, Pages 1928-1941, ISSN 0167-9473, <http://dx.doi.org/10.1016/j.c.sda.2007.06.015>.
  50. Shaban Shataee, Holger Weinaker, Manoucher Babanejad, Plot-level Forest Volume Estimation Using Airborne Laser Scanner and TM Data, Comparison of Boosting and Random Forest Tree Regression Algorithms, *Procedia Environmental Sciences*, Volume 7, 2011, Pages 68-73, ISSN 1878-0296, <http://dx.doi.org/10.1016/j.proenv.2011.07.013>.
  51. Song feng Zheng, QBoost: Predicting quantiles with boosting for regression and binary classification, *Expert Systems with Applications*, Volume 39, Issue 2, 1 February 2012, Pages 1687-1697, ISSN 0957-4174, <http://dx.doi.org/10.1016/j.eswa.2011.06.060>.
  52. L.I. Kuncheva, M. Skurichina, R. P. W. Duin, An experimental study on diversity for bagging and boosting with linear classifiers, *Information Fusion*, Volume 3, Issue 4, December 2002, Pages 245-258, ISSN 1566-2535, [http://dx.doi.org/10.1016/S1566-2535\(02\)00093-3](http://dx.doi.org/10.1016/S1566-2535(02)00093-3).
  53. D.Lu & Q. Weng (2007): A survey of image classification methods and techniques for improving classification performance, *International Journal of Remote Sensing*, 28:5, 823-870, <http://dx.doi.org/10.1080/01431160600746456>.
  54. {R Core Team}, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2013, <http://www.Rproject.org>



This page is intentionally left blank





# Location Based Path Guiding System for the Visually Impaired People

By M. Chowdhury, T. Alam, P.P.Purohit & S.A. Fahim

*Chittagong University of Engineering & Technology, Bangladesh*

**Abstract-** Visually impaired or blind people faced tremendous challenges daily in accessing the information while on the move or during communicating with the world around them. They had to depend on their sighted colleagues, friends, relatives, or the people on the road to help them with locating places. Our Location Based Information System is such a location based application that works as an audio road guider for the visually impaired people using OpenStreetMap. OpenStreetMap (OSM) is a world-wide campaign for developing open source maps. Voice recognition represents a technology that is finally ready for prime time use. In our system we used voice recognition as visually impaired people unable to text searching.

**Keywords:** *visually impaired, openstreetmap, audio direction, android.*

**GJCST-C Classification:** *1.2.10*



LOCATIONBASEDPATHGUIDINGSYSTEMFOTHEVISUALLYIMPAIREDPEOPLE

*Strictly as per the compliance and regulations of:*



RESEARCH | DIVERSITY | ETHICS

# Location based Path Guiding System for the Visually Impaired People

M. Chowdhury<sup>α</sup>, T. Alam<sup>σ</sup>, P.P.Purohit<sup>ρ</sup> & S.A. Fahim<sup>ω</sup>

**Abstract-** Visually impaired or blind people faced tremendous challenges daily in accessing the information while on the move or during communicating with the world around them. They had to depend on their sighted colleagues, friends, relatives, or the people on the road to help them with locating places. Our Location Based Information System is such a location based application that works as an audio road guider for the visually impaired people using OpenStreetMap. OpenStreetMap (OSM) is a world-wide campaign for developing open source maps. Voice recognition represents a technology that is finally ready for prime time use. In our system we used voice recognition as visually impaired people unable to text searching. The CloudMade API was used to provide multi-mode routing services together with turn-by-turn descriptions for car users, bicycle riders, and pedestrians. This solution is relatively easy and fast to deploy. This paper provides a guiding solution with detailed navigational services for visually impaired people. The experimented result demonstrates the effectiveness of our system.

**Keywords:** *visually impaired, openstreetmap, audio direction, android.*

## I. INTRODUCTION

A location-based service (LBS) is an information and entertainment service, accessible with mobile devices through the mobile network and utilizing the ability to make use of the geographical position of the mobile device. Location-based services are a general class of computer program-level services used to include specific controls for location and time data as control features in computer programs. As such (LBS) is an information and has a number of uses in Social Networking today as an entertainment service, which is accessible with mobile devices through the mobile network and which uses information on the geographical position of the mobile device.

For blind people the traditional devices to aid mobility have been the long cane and the guide dog. Although a large number of electronic mobility devices have been developed, their uptake has been very low. To help with orientation, blind people have had to rely on recognizing non-visual landmarks possibly in conjunction with an embossed map of the area.

The blind people faced tremendous challenges daily in accessing the information while on the move or during communicating with the world around them. They

the people on the road to help them with locating places. Considering this, our proposed system is design for the blind people. Most of the existing applications are for normal people and very few of them address the special requirements for the disabled people. About 15 percent of the world's population some 785 million people has a significant physical or mental disability. According to a new report prepared jointly by the World Health Organization [1] and the World Bank [2], Geneva, June, 2012 - 285 million people are visually impaired worldwide: 39 million are blind and 246 have low vision. About 90% of the world's visually impaired live in developing countries. Due to growing populations and aging, these numbers are expected to double by the year 2020 making the colossal human tragedy even worse, stalling development and denying a basic human right [3].

To understand how developing mobile location services applications differs from traditional wire line applications, it is important to understand the basic principles of wireless network [4]. There are two basic of positioning a device on the market today, firstly by using satellite for example, GPS or its Russian cousin, GLONASS; secondly by using the mobile telephone network. There are also combinations of these, which increase the accuracy and decrease the cost for the user.

Global Positioning System (GPS) is based on twenty four satellites and five monitoring stations around the world. It is satellite-based navigation system, which is maintained and operated by the U.S. Department of Defense. Its primary purpose is to provide global positioning and navigation for land-based, sea-based, and air-based tactical forces, but it also been made available for commercial use [4].

GPS works by the user's receiver (which can now be made so small that they can be built into watches and cell phones) acquiring a sufficient number of satellites (a minimum of three) and triangulating them by using the travel time of the radio signal from the satellite (more properly, trilateration or resection because the system is not calculating angles). Because the position of the satellite in a space known (because the orbit and the time is known) and the distance is known (because the time it takes for the signal to get from the satellite is known, with a known velocity) the distance arc on the Earth's surface can be calculated from satellite although delays and errors that might

*Authors α σ ρ ω : Chittagong University of Engineering & Technology  
Department of computer science & engineering, Bangladesh.  
e-mail: mahfuz\_csecuet@yahoo.com*

occur when signal passes through the atmosphere have to be corrected. Intersecting two spheres from satellite gives an ellipsoid; applying the third satellite gives two points where the ellipsoids intersect, the intersection with fourth sphere will narrow it down to one point although usually, one of the two points is not reasonable answer and can be discarded.

Android [5], [6] is an open-source software platform developed by Google, for mobile application development on devices powered by the Android OS. It is a complete software stack that provides all the middleware needed to run end-user applications on mobile devices such as: device drivers, OS, core libraries, an optimized virtual machine, Java Native Interface (JNI), and a complete Java development environment. The blind/visually impaired people usually face the problem in finding location in any area. Hence, we have proposed a location based service on mobile phones using Open Street Map (OSM) which is mainly utilized to provide audio and visual messages on the map [7]-[9]. Open Street Map is a rapidly growing open source map of the world because of the availability of map information across the world and invent of inexpensive portable GPS devices [10]. The demand of location based services is also increasing day by day with the emerging growth of smart phones. Our proposed system is developed for both the normal and blind people. Users of our application will get audio warning message in any direction. The rest of the paper is organized as follows. In Section II, we discuss about some related works and their shortcomings. Then we present the description of proposed location based system in Section III. In Section IV, system implementation details are described. In Section V, we have discussed the experimented result. Finally, we conclude this paper stating the future plan in Section VI.

II. RELATED WORKS

There exists some research works for helping blind people with map technology. Some modifications of these applications are required so that blind people can use these smoothly. An android platform based mobile application using OSM to help visual impaired and normal people was proposed in [11]. This application helps users to search a specific location with text. This system can't give audio direction to reach his/her destination and failed to give audio message if he/she on wrong path. One of the mentionable existing research works was Walky-Talky [12]. This application has been designed to help the blind with walking directions that combine Google Maps with GPS navigation technology. The main weakness of this project was it uses Google Map. Google Map has legal or technical restrictions on its use, holding back people from using it in creative, productive, or unexpected ways. System proposed in [4] and [13] can't consider audio

based path finding help. RFID-based system for navigation in a building for blind or visually impaired people was proposed in [14]. This system can't work for outdoor users.

III. METHODOLOGY

a) Overview of the Overall System

The aim of our work is to help visually impaired people with audio to reach his/her destination. Fig. 1 demonstrates the workflow of our proposed location based system for the visually impaired people.

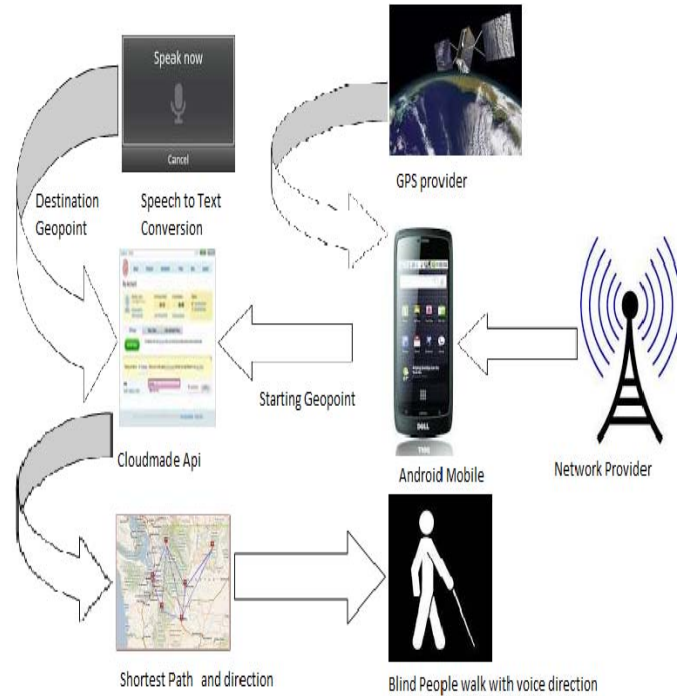


Figure 1 : Proposed location based system for blind people

b) Voice Recognition

First step is taking audio input about a particular location from user and convert the audio to text using "audio to text convert API" of android. The text will generate geopoint of the corresponding location.

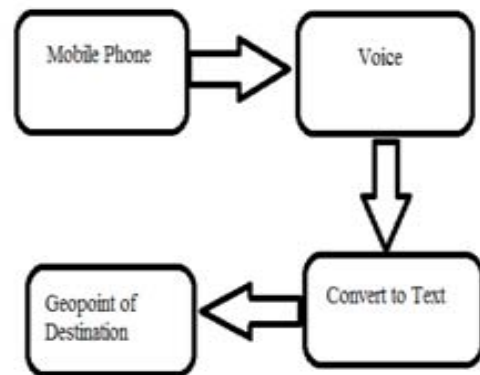


Figure 2 : Conversion from voice to text and assigning geopoint

c) *Cloudmade API*

In this stage GPS provider or network provider will generate the current position of the location and we get destination geopoint from voice recognition. Then the starting and destination geopoints will be generated to Cloudmade API. Cloudmade API will generate geopoints of shortest path between starting and destination geopoints. It also generates direction to reach the destination.

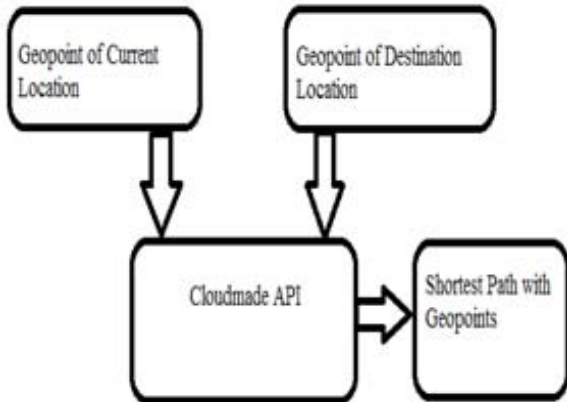


Figure 3 : Shortest path from cloudmade API

d) *Audio Direction*

At this stage user will get audio direction at every turn of road from source to destination. User will get direction on his/her way.

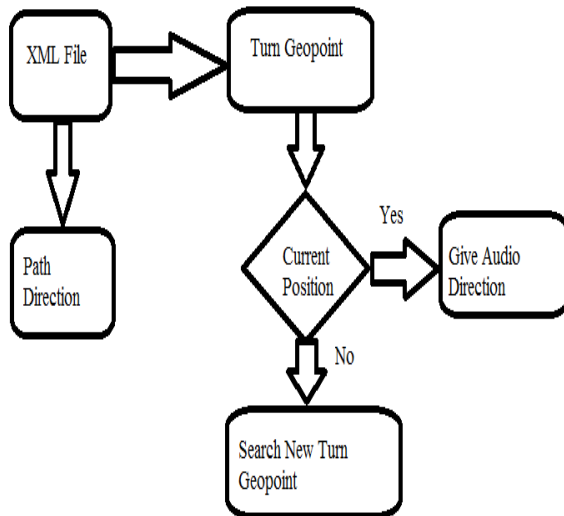


Figure 4 : Audio direction to a specified location

VI. IMPLEMENTATION

a) *Turn on GPS*

The application demands a GPS enabled android mobile. GPS must be enabled before start the application. Network provider is another system which can generate location information to the device from the network operator. In our application we need accurate data of a position. We used GPS provider and Network provider both.



Figure 5 : Turn on GPS to generate location information

b) *Retrieving Voice Recognition Result*

We used Google voice recognition API to convert voice to text. It is a real time service which generates instant text after giving audio to the device.

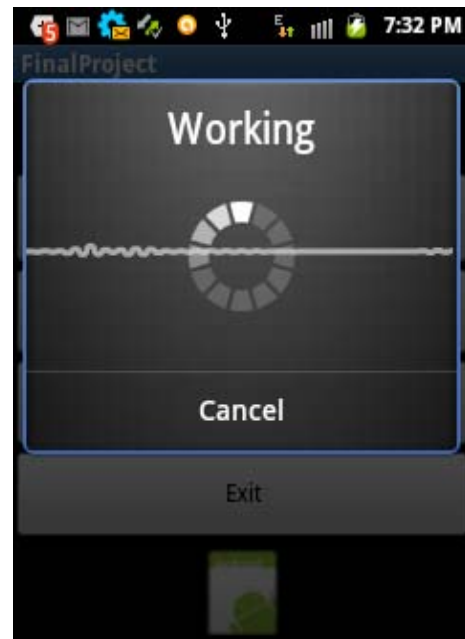


Figure 6 : Voice recognition demo

c) *Retrieve Shortest Path from Cloudmade API*

We have got starting geopoint from network provider and GPS provider and destination geopoint from the user. These geopoints will be generated to the Cloudmade API. There are some formats to generate data to the Cloudmade API. We have selected language English and distance in miles and API 0.3. After giving the geopoints it generates a shortest.xml file. The file contains shortest path from starting location to destination location.

```

<rtept lat="22.462374" lon="91.973289">
<desc>Head south, 98 ft</desc>
<extensions>
<distance>30</distance>
<time>7</time>
<offset>0</offset>
<distance-text>98 ft</distance-text>
<direction>S</direction>
<azimuth>180.0</azimuth>
</extensions>
</rtept>
<rtept lat="22.462105" lon="91.973297">
<desc>Turn right, 0.1 miles</desc>
<extensions>
<distance>216</distance>
<time>52</time>
<offset>2</offset>

```

Figure 7: XML data from cloumade API

#### d) XML Parsing

Parsing XML from the internet is a pretty common requirement on any platform. There are numerous ways to accomplish this task on every language out there, and the Android platform is no exception. On the Android, this is accomplished using SAXParser, a serial access parser API for XML. SAXParser functions as a stream parser, with an event-driven API, using callback methods that trigger every time events occur during the reading. The majority of the work is done by a SAX-Handler. The SAXParser will walk through the XML file from beginning to end (hence parsing is always unidirectional) and calls appropriate handler methods along the way. For this exercise, we will create a Handler that extends org.xml.sax.helpers.DefaultHandler and overrides the necessary methods.

On the start/end of each document, the following methods get called:

```

public void startDocument() throws SAXException {}
public void endDocument() throws SAXException {}

```

When the Parser reaches an opening tag, like <exampletag name=\|abs">, the following method gets called: public void startElement (String namespace URI, String local Name, String qName, Attribute satts) throws SAXException {}

In this case, localName will be \|exampletag". The atts variable will hold any associated attribute information: atts.getValue("name") will return \|abs". When we reach a closing tag, like <=exampletag>, the equivalent closing method gets called: public void endElement(String namespaceURI, String localName, String qName) throws SAXException {}.

In between an opening and closing tag, there can be a string, like <exampletag>here is some content<=exampletag>. The SAXParser reads in the string, one character at a time, but buffers method calls to the handler: public void characters(char ch[], int start, int length) {}.

The ch[] array holds a buffer of characters that the SAXParser has read in, but the only relevant chunk is given by the start and length values. With large enough strings, the characters() method may be called multiple times within a single block of character data. This is a place where we personally stumbled with, as it seems many tutorials out there seem to ignore this fact, assuming the entire block is returned and end up only getting partial data.

#### e) Audio Direction

In the result our application will generate audio direction at every turn of a road. We got geopoints of every turn but we take a round area by taking the geopoint as a center. By Haversine formula [15] we find out the distance between two geopoints and we took 5 meters as its range. We used an algorithm to generate the audio message when user reaches to the area. We used a thread to perform task in background and on progress of the application.



Figure 8: User current position to generate audio direction

## V. EXPERIMENTAL RESULTS AND ANALYSIS

#### a) Experiment

We tested our program by six users. They used our program for 3 days and gave their feedback.

Table 1: Experimented Results

User	Details	Comment
Volunteer-1	Blind Man	Satisfactory
Volunteer-2	Student	Satisfactory
Volunteer-3	Student	Satisfactory
Volunteer-4	Blind Man	Average
Volunteer-5	Blind Man	Average
Volunteer-6	Blind Man	No Comment

### b) Evaluation

After three days of testing period, they acclaimed that the system is much more helpful to find their destination. They found the system much more helpful as they didn't have to search their destination with text typing. Although they were on wrong direction they got new direction from the application. When the volunteer reaches within an area they got audio message. They didn't have to be attentive on their way for searching turn of road.

Criticisms: But still they (users) have some complaints. Volunteer-2 said that it is very slow. We said him, due to slow internet connection it works slowly. They also got some problems in running the application. The application became stop when there is no route in Cloudmade API for given geopoints. We said them about our limitations, in our country OpenStreetMap is not well developed like Google map. Besides, all the routes are not developed to get shortest path and direction of road.

### c) Our Feedback

We have considered all the problems and suggestions from our experiment and fixed the problems. The following solutions of problems are fixed in the new version of our software. Firstly, the speed of speech engine is now maintained so that users can comfortably recognize the audio. Secondly, the distance of two positions is maintained. We hope our new version will now satisfy all the users and perfectly applicable to the real environment.

## VI. CONCLUSION

Our location based system mainly developed for the visually impaired people. Our application helps visually impaired people to reach his/her destination with a shortest path. It generates voice before reaching the turn of road so that user can be aware of the turn of road. There are some limitations for the application to run in developing country. Our total work depends on internet speed. In developing country it is very costly to generate high speed internet. Lack of details on OSM of Bangladesh is the main challenge of our work. The more volunteers working on OSM can make map of the country rich. In that case, our application will perform better in the context of Bangladesh. Moreover, we have future plan to implement another application that will generate suitable path for visually impaired/blind people to reach his/her destination by observing traffic jam and weather. Finally, our proposed system has opened further research so that anyone from any part of the world can work on the existing system and incorporate their ideas with it.

## REFERENCES RÉFÉRENCES REFERENCIAS

1. The World Health Organization website [online]. Available: [http://www.who.int/en/].
2. The World Bank Website [Online]. Available: [http://www.worldbank.org/].

3. BlindPersonStatistics[online].Available:[http://www.who.int/mediacentre/news/releases/pr79].
4. A. Devaraju, "Location Based Service", Und-ergraduate Thesis, September 2004. Available: [http://ifgi.unimuenster.de/~deva01/images/Anusuriya.pdf ].
5. Android[Online]:[http://developer.android.com/guide/basics/what-is-android.html].
6. Androiddeveloper[online].Available:[http://developer.android.com/guide/developing/devices/emulator.html].
7. Bing aerial imagery in OSM[online], Available : [http://wiki.openstreetmap.org/wiki/Bing, retrieved: 9-December-12.
8. ProjectLibya[http://wiki.openstreetmap.org/wiki/WikiProject\_Libya].
9. ProjectHaiti[http://wiki.openstreetmap.org/wiki/WikiProject\_Haiti].
10. OpenStreetMapfoundation[online],Available:[http://www.osmfoundation.org/wiki].
11. Rifat, Moutushy, Ahmed, Ferdous, "Location based Information System using OpenStreetMap", in Research and Development (SCORED), IEEE Student Conference, Dec. 19-20, 2011.
12. Jianghua Zheng, Xiaoyu Chen, Blazej Cierny, Adam C. Winstanley , Peter Mooney and Ricky Jacob, "Mobile routing services for small towns using cloudmade api and OpenStretMap", in The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 38, Part II, pp. 337-342.
13. Moller, Toth, Wang, Moller, Arras, Bach, Schumann and Guttmann, "Enhanced Perception for Visually Impaired People", in Bioinformatics and Biomedical Engineering, 3rd International Conference on Bioinformatics and Biomedical Engineering, 11-13 June, 2009.
14. Chumkamon, P. Tuvaphanthaphiphat, P. keeratiwintakorn, "A Blind Navigation System Using RFID for Indoor Environments", in Electrical Engineering/Electronics, 5th International Conference on Computer, Telecommunications and Information Technology, May 14-17, 2008.
15. Haversineformula[http://en.wikipedia.org/wiki/Haversine\_formula].

This page is intentionally left blank





## A Framework for Context-Aware Semi Supervised Learning

By Vijaya Geeta Dharmavaram & Shashi Mogalla

*GITAM University, India*

**Abstract-** Supervised learning techniques require large number of labeled examples to build a classifier which is often difficult and expensive to collect. Unsupervised learning techniques, even though do not require labeled examples often form clusters regardless of the intended purpose or context. The authors proposes a semi supervised learning framework that leverages the large number of unlabeled examples in addition to limited number of labeled examples to form clusters as per the context. This framework also supports the development of semi supervised classifier based on the proximity of unknown example to the clusters so formed. The authors proposes a new algorithm namely “Semi Supervised Relevance Feature Estimation”, (SFRE), to identify the relevant features along with their significance weightages which is integrated with the proposed framework.

**Keywords:** *context – aware, semi supervised learning, feature relevance, subspace clustering, discriminant analysis.*

**GJCST-C Classification :** *1.2.6*



*Strictly as per the compliance and regulations of:*



# A Framework for Context-Aware Semi Supervised Learning

Vijaya Geeta Dharmavaram <sup>α</sup> & Shashi Mogalla <sup>σ</sup>

**Abstract-** Supervised learning techniques require large number of labeled examples to build a classifier which is often difficult and expensive to collect. Unsupervised learning techniques, even though do not require labeled examples often form clusters regardless of the intended purpose or context. The authors proposes a semi supervised learning framework that leverages the large number of unlabeled examples in addition to limited number of labeled examples to form clusters as per the context. This framework also supports the development of semi supervised classifier based on the proximity of unknown example to the clusters so formed. The authors proposes a new algorithm namely "Semi Supervised Relevance Feature Estimation", (SFRE), to identify the relevant features along with their significance weightages which is integrated with the proposed framework. Experiments conducted on the benchmark datasets from UCI gave results which are very promising and consistent even with lesser number of labeled examples.

**Keywords:** context – aware, semi supervised learning, feature relevance, subspace clustering, discriminant analysis.

## I. INTRODUCTION

Machine learning techniques are being adopted by various applications from different domains to build predictive models. These techniques are broadly classified as supervised learning and unsupervised learning based on the availability of class labels to build the model. Supervised learning methods require labeled data to build a classifier model that predicts the class labels of unknown examples based on the information available in the form of class labels. However, it is usually very expensive and time-consuming process to collect the labeled data (Han et al., 2011). Even in domains with abundance of unlabeled data, labeled data are usually scarce and would require some effort to collect such data. However, to build classifier with better generalized accuracy, large number of labeled data is required, more so for datasets with high dimensionality - one of the problems associated with curse of dimensionality (Ramona et. Al.,2012).

Accordingly, it is believed that with fixed number of labeled examples, the predictive power of the

classifier decreases with the increase in number of dimensions thus requiring larger number of labeled examples for building classifier (Advani, 2011).

In unsupervised learning methods such as clustering, unlabeled data, if available in abundance, suffice to extract hidden patterns of knowledge from a given dataset. Traditional clustering algorithms take into account the entire feature space to partition the datasets into clusters such that there is homogeneity among the instances within a cluster. The proximity between the instances in the cluster is measured in terms of distance function. However, with the increase in dimensions, the distance measures employed in the clustering algorithm becomes insignificant and clusters so produced will be meaningless. Hence clustering will full feature space, especially when the number of dimensions are large, may not produce good clusters.

Finding the subset of feature space to produce meaningful clusters comes under the purview of subspace clustering. Subspace clustering focuses on finding a subset of features or a smaller set of transformed features with an aim to define cluster-able object spaces (Han et al., 2011; Sim et al., 2013). In high dimensional datasets due to exponentially large number of subsets of the feature set, subspace clustering techniques have to eliminate enormous possibilities before identifying the appropriate feature space that contain intrinsically significant clusters (Han et al., 2011). The basic research in subspace clustering falls into unsupervised learning as it tries to identify clusters based on the distribution of objects in various feature sub-spaces irrespective of the class labels of the objects. The clusters thus formed may be meaningful but may not be relevant to the intended purpose or context. For instance, the census data is described in terms of different features like social, economic, education, health, etc.,. However, it needs to be clustered in groups depending on the purpose of the data analysis. Features corresponding to social backwardness and eco-nomic status is used to identify the welfare schemes to be adopted, whereas features corresponding to place of living, commutability, etc., are used to decide the location of new amenities centers. In both the cases, features used and their relative significance will vary with the context or purpose thus requiring the clustering algorithm to give proper emphasis to appropriate features in accordance with the context for which the patterns are extracted. Such clustering is referred to as Context-Aware Subspace Clustering.

**Author α:** Associate Professor, Department of Operations, GITAM Institute of Management, GITAM University.

e-mail: vijayageeta@gitam.edu

**Author σ:** Professor, Department of Computer Science & Systems Engineering, AU College of Engineering, Andhra University.

e-mail: smogalla2000@yahoo.com

Context-aware-subspace clustering aims to find appropriate feature subspace for a given context represented in the form of class labels of a few labeled examples which are consistent with a large collection of unlabeled examples belonging to the same dataset. To the best of our knowledge, not much research was published in support of feature selection algorithms making use of combination of labeled as well as unlabeled examples. Hence semi supervised feature selection algorithms are needed to be developed for formation of context-aware clusters in domains having only limited examples labeled and the rest being left unlabeled.

Semi Supervised Learning which is an integration of supervised and unsupervised learning; makes use of both labeled and unlabeled examples to build a model (Zhu and Goldberg, 2009). Semi supervised learning has two forms namely semi supervised classification and semi supervised clustering. Semi supervised classification uses both labeled and unlabeled data to build the classifier. Using the limited number of labeled data, probable class labels for the unlabeled data is derived which in turn is added to the pool of labeled data thus increasing the number of labeled examples (Han et al., 2011). The basic assumption in this technique is that the similar data will have same class labels (cluster assumption) (Chapelle et al., 2006; Wang et al., 2012). Different methods like self training, co-training, generative probabilistic models, graph based and support vector machines are used for semi supervised classification (Zhu, 2008). In semi supervised clustering, a large set of unlabeled data is accompanied by a small amount of domain knowledge in the form of either class labels or pairwise constraints (must-link and cannot-link) (Girra et al., 2004; Ding et al., 2012). This domain knowledge is used to guide the clustering of unlabeled data so that the intra-cluster similarities are maximized and inter-cluster similarities are minimized and there exist consistency between the partition and the available knowledge (Gao et al., 2006).

Based on the above arguments, authors proposes context-aware semi supervised subspace clustering framework which leverages the domain knowledge in terms of class labels for at least some of the examples (if labeled examples are expensive) in order to estimate the suitability of the features to the intended cluster solution. Proper selection of features and their relative significance is essential in producing context-aware clusters which are probably uni-class clusters. Uni-class clusters contain all or majority of the elements belonging to same class label which is reflected in terms of cluster purity. The clustering framework is further extended to build a classifier which is referred to as semi supervised classifier that requires minimum information for prediction. The authors also

proposes 'Semi Supervised Feature Relevance Estimation', (SFRE), algorithm to estimate the relevant features and their relative significance in terms of weights that define appropriate subspaces for different targets/context. The framework was tested on a few benchmark datasets from UCI repository which has given promising results.

## II. RELATED WORK

Researchers in the past came up with different methods for semi supervised learning. One popular approach is constrained based clustering. Constraint based methods uses pairwise constraints in the form of must-link and cannot-link that guides the clustering process to partition the data in a way that do not violate these constraints (Wagstaff et al., 2001; Basu et al., 2004; Lu and Leen, 2004). Recently Xiong et al., (2014) proposed an iterative based active learning approach to select pairwise constraints for semi supervised clustering. It uses the concept of neighbourhood that contains labeled examples of different clusters based on pairwise constraints. The uncertainty associated with each point's neighbor is resolved through queries. However, repeated clustering is required with growing list of constraints.

Another popular approach for semi supervised clustering is distance based techniques which is based on the cluster assumption. Yin and Hu (2011) proposed semi supervised clustering algorithm using adaptive distance metric learning where clustering and distance metric learning are performed simultaneously. The clustering results are used to learn the distance metric and the data is projected into a low dimensional space such that data separability is maximized. Gao et al., (2006) focused on semi supervised clustering in terms of features rather than examples. It addresses the problem where labeled and unlabeled dataset have different feature set with few common features.

In terms of feature selection, Padmaja et al., (2010) proposed a dimensionality reduction approach that estimates the significance of features based on the fractal dimensions and accordingly selects a subset of features that are essential to capture the characteristics of the dataset. The algorithm detects all types of correlations among features to identify the essential features after eliminating the redundant and irrelevant features. Kernel based feature selection was also explored by a few researchers (Wang, 2008; Ramona et al., 2012). Clustering based feature selection for classification was proposed by Song et al., (2013) where features are clustered based on graph theoretic clustering method.

Research on feature weighting and ranking concentrated more on supervised learning (Eick et al., 2006; Al-Harbi and Rayward-Smith, 2006; Zhao and Qu,

2009). Most of these research studies initially weigh the features by using some random guess or equal weights. These initial weights are then adjusted accordingly. Such approach may take much time to arrive at the final optimum weights if the initial guess is not appropriate.

This paper deals with semi supervised learning methods with wrapper based feature selection method that uses discriminant analysis results to initialize the weights. These weights are adjusted accordingly in a stepwise refinement process using both labeled and unlabeled examples. The proposed framework is used to develop a classifier and a pertinent cluster solution.

### III. CONTEXT-AWARE SEMI SUPERVISED SUBSPACE CLUSTERING FRAMEWORK

Given a dataset with 'l' number of labeled and 'u' number of unlabeled examples such that  $l < u$ , the proposed context-aware semi supervised subspace clustering framework has a four phase architecture as shown in Fig. 1. The phases are explained in the subsequent paragraphs.

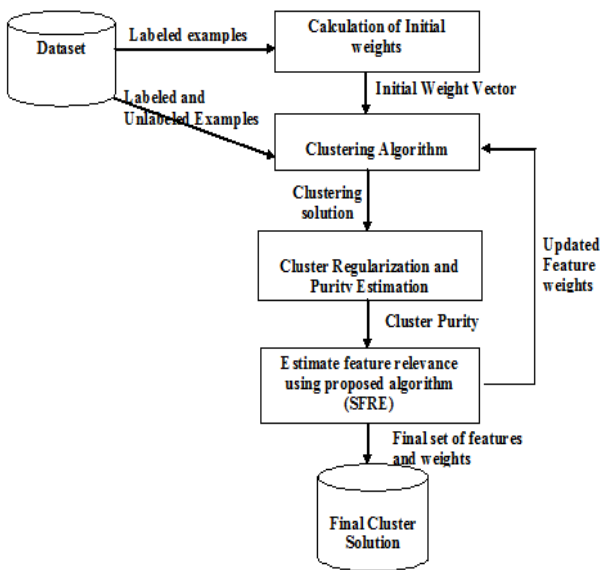


Figure 1 : Architecture of the Proposed Context-Aware Semi Supervised Subspace Clustering Framework

#### a) Initial weights calculation

A dataset may be clustered in multiple ways by appropriately selecting a subset of features /attributes depending on the purpose. Hence to produce clusters conforming to a particular purpose or context, weights must be given to features that depict the importance of the feature. Researchers in the past initially start with a guess/random weights or equal weights to the feature and proceeds further to determine the more acceptable weights. Instead of starting with some arbitrary values, it is proposed to use the information from the available labeled data to initialize the weights which can be adjusted later. Authors thus propose usage of

discriminant analysis that finds the relationship between the independent features (predictors) and the dependent feature (class label), to initialize the feature weights.

Discriminant analysis is a method that is used to predict categorical value from a given set of independent feature. It assumes the independent features to be normally distributed. The linear equation of Discriminant analysis is (Equation 1)

$$D = V_1X_1 + V_2X_2 + V_3X_3 + \dots + V_iX_i + a \quad 1$$

Where

D=Discriminant Score

$V_i$ = the discriminant coefficient or weight of  $i^{th}$  feature

$X_i$ = Value of  $i^{th}$  feature

a = a constant

Discriminant analysis thus identifies the relevant features and its coefficients reflect the relevancy of the feature. The outcome of the discriminant analysis in terms of coefficients is normalized and is used as initial weights for developing binary cluster solution where as development of multi-class cluster solution involves integration of results given through multiple discriminant functions. The proposed framework use potency index as per the approach given in Dharmavaram and Mogalla (2013) for determining the initial weights of various features based on the labeled examples in case of multi-class datasets.

#### b) Clustering Algorithm

The initial weight vector is used to form the initial cluster solution by using any partitional clustering algorithm. The authors have chosen K-means algorithm for its simplicity and computational efficiency to deal with numerical features. While dealing with datasets described in terms of numerical attributes, generally K-means algorithm employs Euclidean distance to compute the distance from each data point to the cluster centroid. Euclidean distance assumes that all the features are equally important while forming the clusters. However, as discussed previously, weights of the feature will determine the relevancy of the feature in forming the desired cluster solution and accordingly Weighted Euclidean Distance metric is used for distance calculation which has the following equation (Equation 2):

$$dw(x_i, x_j) = \sqrt{\sum_{m=1}^p w_m (x_{im} - x_{jm})^2} \quad 2$$

where  $\sum_{m=1}^p w_m = 1$

where  $w_m$  indicates the weight of the  $m^{th}$  feature. If the significance of the feature is more, its weight will be more. The weight of an irrelevant feature can be set to zero.

For clustering, the number of clusters, K, is taken to be more than the number of classes. Larger values of K results in formation of large number of small uni-class clusters and hence, multiple clusters are associated with a single class. Each of these clusters

represents a neighbourhood reflecting the natural grouping or profiling within the same class examples. However, clusters formed with a very few labeled examples may not be enough to define the purity of the clusters. The presence of an outlier labeled example will also alter/hamper the definition of the cluster. Clusters defined in terms of such labeled examples may lead to overfitting that could result in poor predictive performance for unseen examples. The cluster solution thus needs to be regularized to take care of the distribution of the labeled as well as unlabeled examples in the cluster space. The regularization procedure takes into account the presence of unlabeled examples in the vicinity of labeled examples in a cluster while estimating the purity of a cluster. The proximity measure between the labeled and unlabeled examples is used to define the probable class labels of the unlabeled examples. Cluster purity is thus defined in terms of true class labels for the labeled examples and probable class labels for the unlabeled examples.

c) Cluster Regularization and Estimation of Semi supervised Cluster Purity

Given a dataset  $D = L \cup U$  consisting set of 'l' labeled examples, L and 'u' unlabeled examples, U, where  $k < u$ . Each labeled example, n, in L has a label  $y_n \in Q$ , where Q is a set of 'q' number of classes. The author devised a new metric called 'Cluster Concurrence' to measure the purity of the clusters based on the probabilistic class labels of unlabeled examples and true class labels of labeled examples.

In case of labeled examples, true class labels are taken from the dataset. In the case of unlabeled examples, probabilistic class labels  $\hat{y}$ , are generated in terms of the class labels of the labeled examples in their neighbourhood, based on the weighted Euclidean distance between the labeled and the unlabeled examples.

The cluster concurrence is estimated for each cluster based on the agreement of the members of the cluster towards a particular class label and hence reflects the uni-class property of a cluster. In order to estimate the cluster concurrence of  $k^{th}$  cluster, the support,  $S_{kj}$ , available for each class, j, in that cluster is aggregated as shown in Equation 3.

$$S_{kj} = \frac{1}{|C_k|} \sum_{n \in C_k} M \times P_j(n) \quad 3$$

Where

$P_j(n)$  indicates the probability of the example n belonging to the class j

$|C_k|$  is the cardinality of the cluster k, i.e., the number of examples that are assigned to cluster  $C_k$ .

$$M = \begin{cases} 1 & \text{if } j = \underset{i}{\operatorname{argmax}} \{P_i(n)\} \\ 0 & \text{otherwise} \end{cases}$$

The binary term M acts as a deciding factor to indicate whether the example contributes to the support of class j or not. It may be noted that each example, whether labeled or unlabeled, contributes to the support of only one class: the unlabeled example support the class with the maximum probability, while the labeled example naturally support one and only one true class label.

$P_j(n)$  is calculated as per the equation given below (Equation 4) where  $d(i,n)$  is the weighted Euclidean distance between i and n.

$$P_j(n) = \begin{cases} 1 & \text{if } n \in L \text{ and } y_n = j \\ \frac{\sum_{i \in L \cap C_k \text{ and } y_i = j} \frac{1}{d(i,n)^2}}{\sum_{z \in L \cap C_k} \frac{1}{d(z,n)^2}} & \text{if } n \in U \cap C_k \\ 0 & \text{otherwise} \end{cases} \quad 4$$

The predicted label of an unlabeled example, t, is the label for which the probability is maximum.

The cluster concurrence of  $k^{th}$  cluster is estimated as:

$$CC_k = \max_j \{S_{kj}\}$$

Overall cluster purity of the cluster solution is taken as the weighted sum of individual cluster concurrences and is given below (Equation 5)

$$CP = \sum_{t=1}^K \frac{|C_t|}{|D|} CC_t \quad 5$$

d) Estimating feature relevance (SFRE algorithm)

Though discriminant analysis is a proven technique for estimating the relevance of various features in a dataset, for the purpose of semi supervised subspace clustering, it can only make use of the limited labeled examples. In order to estimate the feature relevance more accurately leveraging abundantly available unlabeled examples in the dataset, the author proposes a new algorithm namely Semi supervised Feature Relevance Estimation (SFRE) algorithm, that improves the preliminary estimates made by discriminant analysis and thereby contributes to context-aware semi supervised subspace clustering.

The new algorithm, SFRE is guided by cluster purity estimated in terms of labeled as well as unlabeled examples belonging to various feature subspaces. The algorithm accepts the dataset D that includes L and U, initial cluster purity and the outcome of discriminant analysis as initial weights for formation of initial weight vector as input. The output of the algorithm is accurate relevance estimates of the feature set referred to as weight vector that defines the feature subspace for the given purpose indicated through class labels.

The cluster purity obtained by the initial weights is assigned to current cluster purity as initialization step, after which the algorithm executes the following three steps iteratively:

- Step 1: Finding Relevant Features
- Step 2: Updating Weights
- Step 3: Check for convergence

In the first step, each feature in the feature set is checked for its relevance. Taking one feature at a time, clusters are formed without that feature and cluster purity is estimated. If there is a decrease in cluster purity when compared to the current cluster purity, it indicates that the absence of the feature has resulted in the loss in purity and hence it is marked as relevant feature and its relevance increment is calculated based on the proportionate difference in the cluster purity estimated with and without the feature. If there is increase in the cluster purity when compared to the current cluster purity, it indicates that the absence of the feature has resulted in the gain in purity and hence it is marked as irrelevant feature. The outcome of this step is to mark each feature either relevant or not and to estimate the relevance increment for those relevant features.

In the second step, based on the relevance marking, the weights are adjusted such that weights of the relevant features are incremented in accordance with the relevance increment calculated in step 1. The weights of those features marked irrelevant, are made zero and finally the weight vector is normalized to sum up to 1.

In the final step, clusters are formed with the adjusted weights to judge the final solution. The new cluster purity obtained from clusters formed with updated weights and features is compared with the current cluster purity. If there is improvement in the cluster purity, the new weights are accepted and the new cluster purity is taken as the current cluster purity for comparison in the next iteration. The steps are repeated till there is not much significant improvement in the cluster purity. To change the order in which the features are selected in the subsequent iterations; features are randomly selected without replacement. This supports in avoiding any overlap or correlation in the features and to avoid local maxima.

*e) Formal listing of Proposed Algorithm (SFRE)*

Let  $CP_{curr}$  be the cluster purity estimated for the initial cluster solution then stepwise refinement in weights proceeds as follows:

*Step 1:* For each feature  $x$ , randomly selected without replacement from the feature set  $F$   
 Perform K-means without the feature  $x$  by appropriately normalizing the weight vector  
 Estimate Cluster Purity  $CP_{F-x}$   
 If  $CP_{F-x} < CP_{curr}$  then  
 $x$  is relevant

$$\text{calculate relevance increment, } Rel_x = \frac{CP_{curr} - CP_{F-x}}{CP_{curr}}$$

Else  
 $x$  is not relevant

*Step 2:* Increase the weight of each relevant feature  $x$ ,  
 $W_x = W_x (1 + Rel_x)$   
 For each irrelevant feature  $x$ ,  $W_x = 0$   
 Normalise the weight vector

*Step 3:* Perform K-means with adjusted weights  
 Estimate the cluster purity  $CP_{new}$   
 If  $CP_{new} > CP_{curr}$   
 Accept new weights  
 $CP_{curr} = CP_{new}$   
 Perform above steps till there is no improvement in the cluster purity.

The final cluster solution thus formed consists of context-aware clusters with final set of relevant features and weights.

#### IV. SEMI SUPERVISED CLASSIFICATION FRAMEWORK

The above proposed clustering framework is further extended to support classification. To build the classifier model, the dataset is first divided into two sets: Training Set,  $TR$  and Test Set,  $TS$ . The test set contains only labeled examples as it is required to estimate the error based on the true class labels. However, the training set contains both labeled and unlabeled examples as the semi supervised learning makes use of unlabeled examples based on their vicinity to labeled examples. The model building process has three phases:

*Phase 1: Training the Classifier:* The training set,  $TR$ , is clustered as per the framework proposed in the previous section. The final clusters formed are checked for their decisiveness in predicting the class label. If a cluster is decisive, the cluster is labeled with the majority class label within the cluster otherwise the cluster is not labeled and the details of the cluster is stored in order to apply weighted nearest neighbour classification while classifying unknown examples.

*Phase 2: Testing the Model:* The trained classifier is tested for acceptance by estimating its accuracy in predicting the class labels of the test examples. The predicted class labels are compared with the true class labels to calculate the global error. If the global error is tolerable, the model is finalized otherwise the model is re-trained by changing the number of clusters and the test process is repeated.

*Phase 3: Using the model for Classification:* Once the final classifier is built, class labels of unknown examples are predicted as per the classification rules of the model.

*a) Estimating the decisiveness of a cluster*

A cluster which is considered as a uni-class cluster has most of its members agreeing on a particular class. Such cluster also referred to as decisive cluster, is labeled with the (majority) class label which has the maximum support of the examples in that cluster.

However, in the presence of overlapping examples or outliers, the examples in a cluster may not strongly agree on a particular class and such cluster is not considered as uni-class / decisive cluster and is not labeled as they are considered as indecisive cluster. The final cluster solution formed in the training phase contains K clusters with each cluster containing examples belonging to one or more classes. The support of a class in a cluster  $S_{k_j}$  is estimated in terms of true class labels of labeled examples and the predicted (probabilistic) class labels of unlabeled examples in the  $k^{\text{th}}$  cluster. In a given cluster, the difference between the support available for majority class and its competing class reflects the decisiveness of the cluster in concurrence with the majority class. For this purpose, the authors propose a metric referred to as 'Purity Margin' which is measured for each cluster and is compared against purity threshold as detailed below.

#### b) Purity Threshold of the cluster

The 'Purity Threshold',  $PT_k$ , of a cluster,  $C_k$ ,  $PT_k$  is set as the minimum difference or margin, to be imposed between two competing classes in a cluster, for it to be considered as the decisive cluster. The purity threshold is estimated as a pre-defined fraction ( $\lambda$ ) of the product of cluster concurrence  $CC_k$  and the number of classes in the dataset. In a dataset with  $q$  classes, purity threshold  $PT_k$ , for a cluster  $C_k$  is calculated as (Equation 6)

$$PT_k = \lambda \cdot CC_k \cdot q \quad 6$$

Various experiments conducted on the value of  $\lambda$  shows that 0.1 which indicates 10% of support value, is a good measure to get optimum purity threshold.

#### c) Purity Margin of the cluster

The purity margin measures the difference between the maximum support of a class in a cluster and the support of its immediate competitor class. Larger the margin, more pure the cluster is. Intuitively it is taken that it should be greater than or equal to the purity threshold.

For a cluster  $C_k$ , the purity margin  $PM(C_k)$  is calculated as (Equation 7)

$$PM(C_k) = CC_k - S_{k_p} \text{ where } p \text{ is the competing class.} \quad 7$$

#### d) Decisive cluster

A cluster  $C_k$  is considered to be a uni-class or a decisive cluster, if  $PM(C_k) \geq PT_k$  else it is considered as indecisive cluster. The decisive cluster is labeled with the majority class label i.e., the class label that has maximum support of the examples in the cluster, over all classes in the cluster. The indecisive cluster is left unlabeled and the details of the cluster including the predicted labels of unlabeled examples are stored to apply the weighted nearest neighbour classification while classifying any unknown / test example.

#### e) Hybrid Model for Classification

The authors propose a hybridization of model-based classification and instance-based classification for classifying any unknown / test example based on whether it is compatible to decisive cluster or an indecisive cluster.

Let the cluster,  $C_k$  be the most compatible cluster for unknown example  $x$ :

- If the cluster,  $C_k$  is decisive then
- Assign the cluster label,  $y_{C_k}$  to the example  $x$ .
- If the cluster is indecisive then
- Apply weighted nearest neighbor classification to predict the class label of  $x$ .

#### f) Finding the most compatible cluster for unknown / test example

Consider a set of clusters  $C = \{C_1, C_2, \dots, C_K\}$  with centroids as  $c = \{c_1, c_2, \dots, c_K\}$ . Weighted Euclidean distances are calculated between unknown / test example,  $t$ , and each centroid,  $c_i$ . The cluster  $C_k$ , which has the minimum distance among all the clusters is said to be the most compatible cluster for the example,  $t$ . Mathematically, it may be expressed as (Equation 8)

$$k = \underset{i}{\operatorname{argmin}} \{d(t, c_i)\} \quad 8$$

Hybrid model for classification is applied on the value of  $k$  as discussed earlier.

#### g) Weighted Nearest Neighbour Classification

If the cluster  $C_k$  is an indecisive cluster, the cluster contains examples from more than one class. In other words, the cluster contains multiple neighborhoods dominated by different class labels. Hence, the proximity of the unknown / test example,  $t$ , to each class must be measured. The closer the example,  $t$ , is to the neighborhood dominated by particular class label, it is more likely to share the same class label of its neighbors (Cluster Assumption). Accordingly, all the members of the most compatible cluster  $C_k$  are considered as neighbors with weights assigned in the inverse proportion of their squared distance to the test example. The proximity of the example,  $t$ , to a class label,  $p$ , denoted by  $W_{tp}$ , is estimated by aggregating the weights of the members belonging to that particular class. Mathematically it may be expressed as (Equation 9)

$$W_{tp} = \sum_{i \in C_k \text{ and } \hat{y}_i = p} \frac{1}{d(t, i)^2} \quad 9$$

where  $d(t, i)$  is the Euclidean distance between  $t$  and  $i$ .

This proximity estimate will ensure that the examples that are far (possibly an outlier) from the test example has less impact on prediction compared to the ones that are closer by.

The unknown / test example is assigned the class label for which the proximity is maximum (Equation 10).

$$\hat{y}_t = \underset{p}{\operatorname{argmax}} (W_{tp}) \forall p \in Q_k, \text{ the set of classes in the cluster } C_k. \quad 10$$

## V. EXPERIMENTS AND RESULTS

### a) Experimental Setup

The proposed model was implemented on Intel Pentium dual core processor with 3GB of DDR2 667 MHz memory and coded using .NET framework. SPSS statistic tool is used for performing discriminant analysis.

Experiments were conducted on benchmark datasets obtained from UCI repository and one dataset from SPSS Inc. to test the performance of the proposed framework. Five binary datasets and six multi-class datasets were used in the experimentation as shown in table 1.

Table 1 : Description of Datasets

S.No.	Dataset	#Instances	# Attributes	Class
1.	Breast Cancer	683	9	2
2.	Credit	690	15	2
3.	Ionosphere	351	34	2
4.	Pima	768	8	2
5.	Bankloan	700	8	2
6.	Ecoli	336	7	8
7.	Glass	214	9	10
8.	Iris	150	4	3
9.	Wine	178	13	3
10.	Yeast	1484	8	10
11.	Zoo	101	7	7

The labels from some of the examples were purposefully removed to consider them as unlabeled examples and the percentage of labeled examples were varied from 75% to 15% to observe the change in the model's performance with decreasing percentage of labeled examples.

In case of semi-supervised classification the dataset is split into training dataset and test dataset in the ratio of 75:25. Class labels were removed from appropriate number of examples in the training set to test the suitability of the model as a semi supervised classifier. The labeled and unlabeled examples in the training set are taken randomly in varied percentages, in the ratio of 75:25, 50:50, 25:75 and 15:85. The proposed model is trained on the training dataset and the accuracy of the model is estimated based on the test dataset by comparing the true class labels of test dataset with the predicted class labels of the model.

For binary class datasets, experiments were conducted with 100% labeled examples to assess the performance of the framework when all the examples in the datasets are labeled. However availability of labeled examples upto 100% does not call for semi supervised learning. The case with 100% labeled examples was demonstrated only to prove that the proposed method

can handle datasets having less labeled examples in the similar way with datasets having 100% labeled maintaining consistently high performance. The complexity of cluster regularization and estimation of cluster concurrence and purity margin for development of hybrid classifier are not required for datasets having near 100% labeled examples and they may be better processed by an appropriate supervised learning algorithm. The performance of the model for multi-class datasets was analysed starting from 75%.

In both the cases of clustering and classification, discriminant analysis is performed using SPSS statistics tool on the labeled examples in the datasets to produce the discriminant function(s). For binary class datasets, discriminant coefficients, and for multi-class datasets, potency index values are used to get the initial weights of the features in the dataset, which are referred to as initial weight vector.

### b) Results

In case of Semi Supervised Subspace Clustering, the cluster purity was estimated based on the cluster concurrence and the number of relevant features identified for the benchmark datasets are tabulated in table 2 and 3. The change in cluster purity with varied percentage of labeled examples is shown in figure 2 and figure 3.

Table 2 : Cluster Purity of Context-Aware Semi Supervised Subspace clustering – Binary Class Datasets

S.No	Dataset	100%	75 %	50 %	25 %	15 %
1	Bcancer	97.24	96.94	95.76	96.34	96.29
2	Credit	86.52	86.26	85.63	85.77	85.78
3	Ionosphere	90.56	88.23	90.21	88.24	88.56
4	Pima	77.65	76.14	75.86	76.79	77.90
5	Bankloan	80.0	77.92	76.91	77.39	73.94

Table 3 : Cluster Purity of Context-Aware Semi Supervised Subspace clustering – Binary Class Datasets

S.No	Dataset	75 %	50 %	25 %	15 %
1	Ecoli	86.24	82.90	83.82	82.81
2	Glass	72.31	73.72	72.76	69.01
3	Iris	96.64	96.64	95.30	95.92
4	Wine	96.61	97.74	96.61	95.44
5	Yeast	58.04	57.90	56.30	56.10
6	Zoo	84.81	97.0	92.0	65*

\*The size of the zoo dataset is 101. As 15% of the examples could not cover all the seven classes, the error has increased unnaturally.



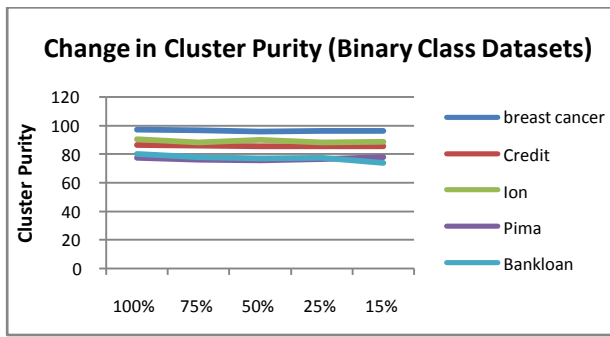


Figure 2 : Changes in cluster purity for Binary Class Dataset

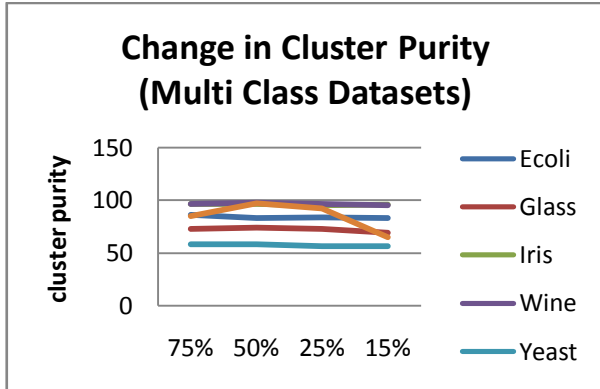


Figure 3 : Changes in cluster purity for Multi Class Dataset

From Fig. 2 and Fig. 3, it is observed that the proposed model has consistent performance in term of cluster purity and not much change is observed with variation in percentage of labeled example. Only in the case of Zoo dataset, there has been huge decline in the cluster purity when there are few labeled examples. This is attributed to the fact, that number of examples in zoo dataset are only 101 and 15% of labeled data is very less compared to number of class labels and may not capture representatives from all the 7 class examples.

In case of Semi Supervised Classification, the training sets of benchmark datasets are used to build the classifier and the accuracy of the classifier is tested on the test set where the predicted class labels are compared with true class labels of the test examples. These test results given in terms of accuracy is compared with the proven classifier models. The models considered for comparison are Weka implementation of C4.5 and an ensemble method, Bagging. Only one ensemble method is considered for comparison as all the other ensemble methods has similar performance on most of the datasets (Tan et al., 2006: Table 5.5). The results are tabulated in table 4 and 5 and a sample comparison graphs for a dataset in binary and multiple class is shown in Fig. 4 and Fig.5.

Table 4 : Comparison table for Semi Supervised Classification – Binary Class Datasets

Dataset	Ensemble – Bagging					C4.5					Proposed Model				
	100	75	50	25	15	100	75	50	25	15	100	75	50	25	15
Breast Cancer	97.56	95.21	95.20	95.09	86.82	94.84	95.24	94.03	91.70	91.61	<b>97.60</b>	<b>97.56</b>	<b>96.68</b>	<b>95.74</b>	<b>95.74</b>
Credit	<b>92.02</b>	81.08	80.41	79.02	77.20	86.37	83.78	<b>80.47</b>	80.0	79.0	86.52	<b>85.21</b>	80.28	<b>80.92</b>	<b>79.51</b>
Ionosphere	94.01	89.47	88.31	86.84	80.51	<b>99.0</b>	<b>92.20</b>	<b>90.78</b>	<b>88.31</b>	<b>84.21</b>	91.76	88.0	86.6	86.64	84.0
Pima	<b>88.93</b>	76.53	74.86	<b>75.69</b>	<b>71.80</b>	84.11	71.82	71.50	70.94	70.39	77.77	<b>76.83</b>	<b>76.27</b>	75.0	70.05
Bankloan	<b>85.23</b>	<b>76.40</b>	74.0	72.0	72.0	90.0	73.93	72.34	72.0	70.0	78.11	74.63	<b>74.62</b>	<b>74.09</b>	<b>72.27</b>

Table 5 : Comparison table for Semi Supervised Classification – Multi-class Datasets

Dataset	Ensemble – Bagging				C4.5				Proposed Model			
	75	50	25	15	75	50	25	15	75	50	25	15
Ecoli	74.66	74.02	70.66	56	76	<b>75.32</b>	70.66	54.66	<b>76.5</b>	75.3	<b>73.9</b>	<b>69.86</b>
Glass	61.66	<b>62.71</b>	49.15	49.15	<b>62.71</b>	61.66	49.15	45.76	60.3	57.72	<b>56.8</b>	<b>57.7</b>
Iris	<b>100</b>	94.11	94.11	58.82	97.11	94.11	91.17	76.47	96.96	<b>96.96</b>	<b>96.5</b>	<b>93.93</b>
Wine	91.66	91.66	88.88	86.11	91.66	91.66	88.88	83.33	<b>97.14</b>	<b>94.2</b>	<b>94.2</b>	<b>91.4</b>
Yeast	<b>58.71</b>	54.15	51.87	45.6	52.9	52.53	51.44	51.74	55.52	<b>54.71</b>	<b>52.5</b>	<b>51.21</b>
Zoo	82.14	77.77	75	53.57	82.14	78.57	77.77	64.28	88.88	<b>85.18</b>	<b>81.48</b>	<b>72.22</b>

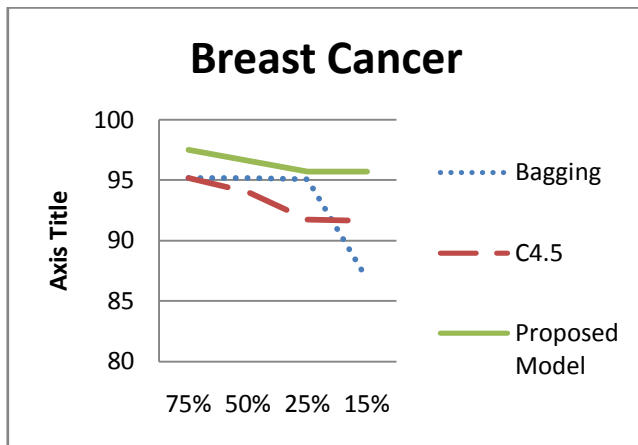


Figure 4 : Comparison Graphs for Breast Cancer (binary class)

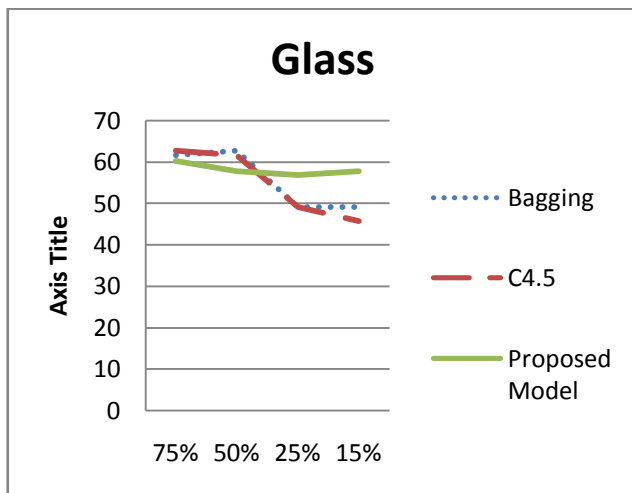


Figure 5 : Comparison Graphs for Glass (multi-class)

It is observed from the comparison tables and graphs that the proposed model has performed consistently. For fewer labeled data (15%), the model performed better than other classifier models. The performance of the model does not degrade much with decreasing number of labeled examples.

Experiments on the benchmark datasets shows that the proposed framework for both clustering and classification have performed consistently better for building models on the training set with varied range (75% to 15%) of labeled examples. When compared to other proven techniques, the proposed framework sustained its performance even when the number of labeled examples is reduced to 15% thus establishing its validity as a semi supervised learning model. The proposed framework was able to identify the relevant features along with their weightages thus reducing the information requirement for handling unknown situations may it be classification or clustering.

## VI. CONCLUSION

In this paper, the authors proposed a framework for context-aware semi supervised learning in terms of both clustering and classification. The proposed framework is useful to work in the domains where availability of labeled data is either scarce or difficult/expensive to obtain. The framework with wrapper based feature selection is very much useful in handling high dimensional datasets. With dimensions reduced, a cluster and classification solution is defined with lesser number of features. This is very useful in cases where there are time and space constraints. The proposed framework not only identifies the relevant features but also estimates the importance of a feature in terms of weights such that cluster solutions are formed as per the intended purpose. Though the framework has used K-means for the formation of cluster solution, the proposed SFRE algorithm can be wrapped into any partitioning clustering algorithm with equal ease for producing context-aware semi supervised subspace clusters leveraging a few labeled examples for defining the context.

Since the model uses discriminant analysis for identifying attributes, it is limited to the numerical data. However, in reality, many of the applications contains mixed data, a combination of numeric and categorical data. This opens an avenue for further research to extend the model to work with categorical data.

## REFERENCES RÉFÉRENCES REFERENCIAS

1. Advani, M. (2011). Learning from High Dimensional fMRI Data using Random Projections. Retrieved from <http://cs229.stanford.edu/>.
2. Al-Harbi, S. H., & Rayward-Smith, V. J. (2006). Adapting k-means for supervised clustering. *Applied Intelligence*, 24(3), 219-226.
3. Basu, S., Bilenko, M., & Mooney, R. J. (2004, August). A probabilistic framework for semi-supervised clustering. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 59-68). ACM.
4. Chapelle, O., Schölkopf, B., and Zien, A. (20-6) *Semi-supervised learning*. Cambridge, MA: MIT press.
5. Dharmavaram, V. G., & Mogalla, S. (2013). Semi Supervised Weighted K-Means Clustering for Multi Class Data Classification. *IJCAIT*, 2(1), 36-44.
6. Ding, S., Qi, B., Jia, H., Zhu, H., & Zhang, L. (2012). Research of semi-supervised spectral clustering based on constraints expansion. *Neural Computing and Applications*, 1-6.
7. Eick, C. F., Rouhana, A., Bagherjeiran, A., & Vilalta, R. (2006). Using clustering to learn distance functions for supervised similarity assessment. *Engineering Applications of Artificial Intelligence*, 19(4), 395-401.

8. Gao, J., Tan, P. N., & Cheng, H. (2006, April). Semi-Supervised Clustering with Partial Background Information. In *SDM*.
9. Grira, N., Crucianu, M., & Boujemaa, N. (2004). Unsupervised and semi-supervised clustering: a brief survey. A review of machine learning techniques for processing multimedia content, Report of the MUSCLE European Network of Excellence (FP6).
10. Han, J., Kamber, M., & Pei, J. (2006). *Data mining: concepts and techniques*. New Delhi: Morgan kaufmann.
11. Lu, Z., & Leen, T. K. (2004). Semi-supervised learning with penalized probabilistic clustering. In *Advances in neural information processing systems* (pp. 849-856).
12. Padmaja P., Shashi M., Teja K.N.K. (2010). Intrinsic Dimensionality Based Conceptual Clustering, *International Journal of Advanced Computer Engineering*, 3(1), 1-5.
13. Ramona, M., Richard, G., & David, B. (2012). Multiclass feature selection with kernel Gram-matrix based criteria. *Neural Networks and Learning Systems, IEEE Transactions on*, 23(10), 1611 – 1623.
14. Sim, K., Gopalkrishnan, V., Zimek, A., & Cong, G. (2013). A survey on enhanced subspace clustering. *Data mining and knowledge discovery*, 26(2), 332-397.
15. Song, Q., Ni, J., & Wang, G. (2013). A fast clustering-based feature subset selection algorithm for high dimensional data. *Knowledge and Data Engineering, IEEE Transactions on*, 25(1), 1-14.
16. Tan Pang-Ning, Steinbach Michael and Kumar Vipin. (2006). *Introduction to Data Mining*. New Delhi: Pearson Education.
17. Wagstaff, K., Cardie, C., Rogers, S., & Schrödl, S. (2001, June). Constrained k-means clustering with background knowledge. In *ICML (Vol. 1, pp. 577-584)*.
18. Wang, L. (2008). Feature selection with kernel class separability. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(9), 1534-1546.
19. Wang, Y., Chen, S., & Zhou, Z. H. (2012). New semi-supervised classification method based on modified cluster assumption. *Neural Networks and Learning Systems, IEEE Transactions on*, 23(5), 689-702.
20. Xiong, S., Azimi, J., & Fern, X. (2014). Active learning of constraints for semi-supervised clustering. *Knowledge and Data Engineering, IEEE Transaction on*, 26(1), 43-54
21. Yin, X., & Hu, E. (2011). Distance metric learning guided adaptive subspace semi-supervised clustering. *Frontiers of computer science in china*, 5(1), 100-108.
22. Zhao, Q., & Qu, H. (2009, June). A Novel Supervised Clustering Based on the Feature Classification Weight. In *Computational Intelligence and Natural Computing, 2009. CINC'09. International Conference on (Vol. 1, pp. 117-120)*. IEEE.
23. Zhu, X. (2006). Semi-supervised learning literature survey. *Computer Science, University of Wisconsin Madison*, 2, 3.
24. Zhu, X., & Goldberg, A. B. (2009). Introduction to semi-supervised learning. *Synthesis lectures on artificial intelligence and machine learning*, 3(1), 1-130.



GLOBAL JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY: C  
SOFTWARE & DATA ENGINEERING

Volume 14 Issue 1 Version 1.0 Year 2014

Type: Double Blind Peer Reviewed International Research Journal

Publisher: Global Journals Inc. (USA)

Online ISSN: 0975-4172 & Print ISSN: 0975-4350

# Supervised Classification of Remote Sensed data Using Support Vector Machine

By Tarun Rao & T.V. Rajinikanth

*Acharya Nagarjuna University, India*

*Abstract-* Support vector machines have been used as a classification method in various domains including and not restricted to species distribution and land cover detection. Support vector machines offer many key advantages like its capacity to handle huge feature spaces and its flexibility in selecting a similarity function. In this paper the support vector machine classification method is applied to remote sensed data. Two different formats of remote sensed data is considered for the same. The first format is a comma separated value format wherein a classification model is developed to predict whether a specific bird species belongs to Darjeeling area or any other region. The second format used is raster format which contains image of Andhra Pradesh state in India.

*Keywords:* classification, data mining, support vector machine, remote sensed data.

*GJCST-C Classification:* G.4



*Strictly as per the compliance and regulations of:*



© 2014. Tarun Rao & T.V. Rajinikanth. This is a research/review paper, distributed under the terms of the Creative Commons Attribution-Noncommercial 3.0 Unported License (<http://creativecommons.org/licenses/by-nc/3.0/>), permitting all non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

# Supervised Classification of Remote Sensed Data using Support Vector Machine

Tarun Rao <sup>α</sup> & T.V. Rajinikanth <sup>σ</sup>

**Abstract-** Support vector machines have been used as a classification method in various domains including and not restricted to species distribution and land cover detection. Support vector machines offer many key advantages like its capacity to handle huge feature spaces and its flexibility in selecting a similarity function. In this paper the support vector machine classification method is applied to remote sensed data. Two different formats of remote sensed data is considered for the same. The first format is a comma separated value format wherein a classification model is developed to predict whether a specific bird species belongs to Darjeeling area or any other region. The second format used is raster format which contains image of Andhra Pradesh state in India. Support vector machine classification method is used herein to classify the raster image into categories. One category represents land and the other water wherein green color is used to represent land and light blue color is used to represent water. Later the classifier is evaluated using kappa statistics and accuracy parameters.

**Keywords:** *classification, data mining, support vector machine, remote sensed data.*

## I. INTRODUCTION

Data mining is the process of extracting useful information from various data repositories wherein data might be present in different formats in heterogeneous environments[1][2]. Various methods like classification, association, clustering, regression, characterization, outlier analysis can be used to mine the necessary information. In this paper we shall be focusing on classification.

Classification is the process wherein a class label is assigned to unlabeled data vectors. Classification can be further categorized as supervised and unsupervised classification. In supervised classification the class labels or categories into which the data sets need to be classified into is known in advance. In unsupervised classification the class label is not known in advance[3]. Unsupervised classification is also known as clustering. Supervised classification can be subdivided into non-parametric and parametric classification. Parametric classifier method is dependent on the probability distribution of each class. Non parametric classifiers are used when the density function is not known[4].

One of the very prominent parametric supervised classification methods is support vector machines(SVM).

In this paper SVM are used to perform the said classification. Herein the data vectors are represented in a feature space. Later a hyperplane that geometrically resembles a slope line is constructed in the feature space which divides the space comprising of data vectors into two regions such that the data items get classified under two different class labels corresponding to the two different regions[5]. It helps in solving equally two class and multi class classification problem[6][7]. The aim of the said hyper plane is to maximize its distance from the adjoining data points in the two regions. Moreover, SVM's do not have an additional overhead of feature extraction since it is part of its own architecture. Latest research has proved that SVM classifiers provide better classification results when one uses spatial data sets as compared to other classification algorithms like Bayesian method, neural networks and k-nearest neighbors classification methods[8][9].

SVM have been used to classify data in various domains like land cover classification[10], species distribution[11], medical binary classification[9], fault diagnosis[12], character classification[5], speech recognition[13], radar signal processing[14], habitat prediction etc... In this paper SVM is used to classify remote sensed data sets. Two formats of remote sensed data viz. raster format and comma separated value(CSV) file formats have been used for performing the said classification using SVM.

Our next section describes Background Knowledge about SVM classifiers. In section 3 materials and methods viz. data acquired and the proposed methodology have been discussed. Performance analysis is discussed in Section 4. Section 5 concludes this work and later acknowledgement is given to the data source followed by references.

## II. BACKGROUND KNOWLEDGE

### a) Overview of SVM Classifier

Support vector machine (SVM) is a powerful tool used in solving either two class or multi class classification problem[15][16]. In a two class problem the input data has to be separated into two different categories wherein each category is assigned a unique class label[17]. A multi class classification problem can be solved by dividing it into multiple two class classification problems and later aggregating the individual results to get the final result of the multi class problem.

Author <sup>α</sup>: Acharya Nagarjuna University, Guntur, India.  
e-mail: tarun636@gmail.com

Author <sup>σ</sup>: Sreenidhi Institute of Science and Technology, Hyderabad, India. e-mail: rajinitv@gmail.com

SVM can be categorized into non-linear and linear SVM. Data can be represented in space as shown in Fig 1. Linear SVM can be geometrically represented by a line which divides the data space into two different regions thus resulting in classifying the said data which can be assigned two class labels corresponding to the two regions[18][19][20].

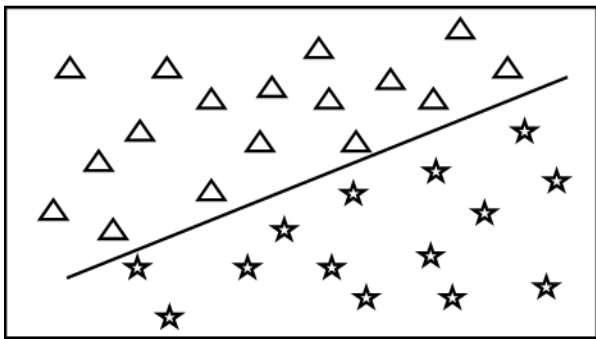


Figure 1 : The Hyperplane

The line mentioned herein is called a hyperplane and can be mathematically represented by equation (1)[21]:

$$\begin{aligned} mx_i + b &\geq +1 \\ mx_i + b &\leq -1 \end{aligned} \quad (1)$$

The data points can be represented by equation (2) [22]:

$$f(x) = \text{sgn}(mx + b) \quad (2)$$

where  $\text{sgn}()$  is known as a sign function, which is mathematically represented by the following equation:

$$\text{sgn}(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases} \quad (3)$$

There can be many hyperplanes which can divide the data space into two regions but the one that increases the distance amid the bordering data points in the input data space is the result to the two class problem. The adjoining data points close to this hyperplane are called support vectors. This concept can be illustrated geometrically as in Figure 2.

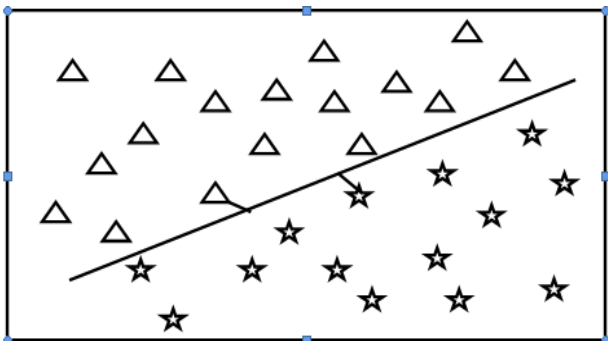


Figure 2 : Distance of the nearest data vectors from the Hyperplane

The margin width can be represented mathematically by the equation:

$$M = \frac{(x^+ - x^-) \cdot m}{|m|} = \frac{2}{|m|} \quad (4)$$

This maximization problem viz. maximizing the distance between the hyperplane and the adjoining support vectors can be represented as a Quadratic Optimization Problem as in equation(5)[22][23]:

$$h(m) = \frac{1}{2} m^2 m \quad (5)$$

subject to  $y_i(mx_i + b) \geq 1, \forall i$

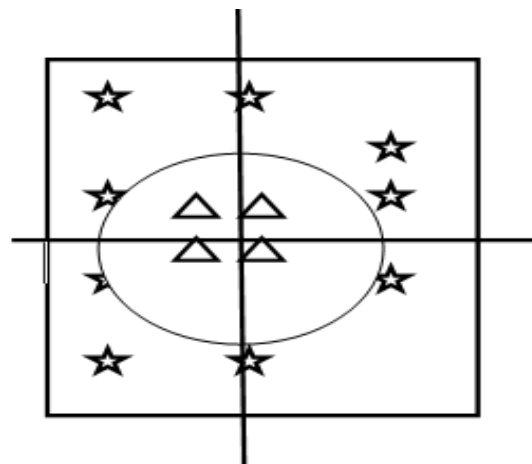
The solution for this problem can be provide by a Lagrange multiplier  $\alpha_i$  which is associated with every constraint in the main problem. The solution can be represented as:

$$\begin{aligned} m &= \sum \alpha_i y_i x_i \\ b &= y_k - m^* x_k \text{ for any } x_k \text{ such that Lagrange multiplier } \alpha_k \neq 0 \end{aligned} \quad (6)$$

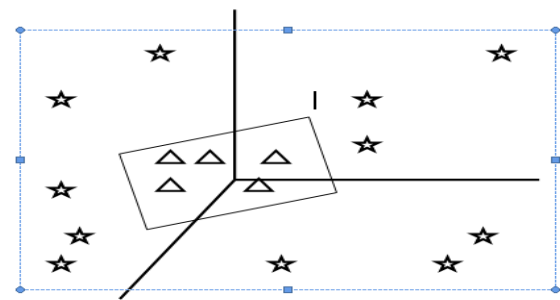
The classifier can be denoted as[16]:

$$f(x) = \sum \alpha_i y_i x_i x + b \quad (7)$$

In the case of non-linear SVM's the input data space can be generalized onto a higher dimensional feature space as illustrated in Fig 3.



(a)



(b)

Figure 3 : (a) Input space (b) Higher dimensional feature space

If every data point in the input data space is generalized onto a higher dimensional feature space which can be represented as [18]:

$$K(x_i, x_j) = f(x_i)^t f(x_j) \quad (8)$$

This is also called a kernel function. It is computed using an inner dot product in the feature space. Various kernel functions can be used to do the said mapping as mentioned in the below equations[23]:

Linear Kernel function =  $x_i^t x_j$

Polynomial kernel function =  $(1 + x_i^t x_j)^p$

Gaussian radial based kernel function =  $\exp(-\frac{|x_i - x_j|^2}{2\sigma^2})$

Sigmoid kernel function =  $\tanh(\beta_0 x_i x_j + \beta_1)$  (9)

One of the major advantages of SVM is that feature selection is automatically taken care by it and one need not separately derive features.

### III. MATERIALS AND METHODS

#### a) Data Acquisition

In this paper SVM classification methodology is applied on two different data set formats. The first format of data sets used is a comma separated value(CSV) file which shall have all relevant attributes necessary for the said classification separated by comma. The data sets used in this category is taken from the birds species occurrences of North-east India[24]. The second format of data sets for classification is in raster format[25]. Raster image is a collection of pixels represented in a matrix form. Raster images can be stored in varying formats. The raster format used herein is TIFF format. A map of Andhra Pradesh state in India used.

#### b) Proposed Method

The data under consideration is first pre-processed. [26]. In the case of csv datasets comprising of information of birds of North-east India the attributes considered are id, family, genus, specific\_ epit-het, latitude, longitude, ver-b-atim \_scientific\_ name, ve-rba-tim\_ family, verbatim\_ genus, verbatim\_ specific\_ ep-ithet and locality. A variable called churn acts as a class label which would categorize the data into two cate-goriesviz onehaving data sets of birds from Darjeeling area and the other having data sets of birds belonging to other north eastern parts in India. Before applying the clas-sification the data sets are cleaned to remove any mis-sing values. In the case of raster data set, a TIFF image is used. The image comprises of a map of Andhra Pradesh, a state in India. Initially a region of interest(ROI) is captured and later supervised SVM classification methodology is applied. Algorithm that explains imple-mentation of SVM is given below [27]:

Begin

Step 1: Loop the n data items

Step 2: Start dividing the input data set into two sets

of data corresponding to two different categories

Step 3: If a data item is not assigned any of the regions mentioned then add it to set of support vectors V

Step 4: end loop

End

### IV. PERFORMANCE ANALYSIS

#### a) Environment Setting

A total of 695 data set records act as test data set and are used to authenticate the classification results obtained for CSV data sets and in the case of TIFF raster data sets one Region of interest is extracted from the given input image. The proposed method has been implemented under the environment setting as shown in Table 1[28][29].

Table 1 : Environment Setting

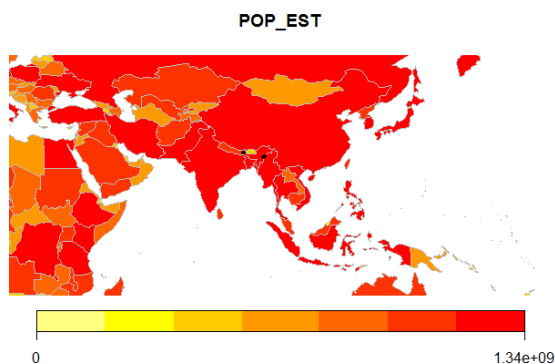
Item	Capacity
CPU	Intel CPU G645 @2.9 GHz processor
Memory	8GB RAM
OS	Windows 7 64-bit
Tools	R, R Studio, Monteverdi tool

#### b) Result Analysis

Classification accuracy can be measured using parameters of a confusion or error matrix view depending on whether the event is correctly classified or no event is correctly classified as shown in Table 2[9]. And the classified results for CSV format data sets is demonstrated in Figure 4.

Table 2 : Confusion / Error Matrix View

Real group	Classification result	
	No Event	Event
No Event	True Negative(TN)	False Positive(FP)
Event	False Negative(FN)	True Positive(TP)



(a)

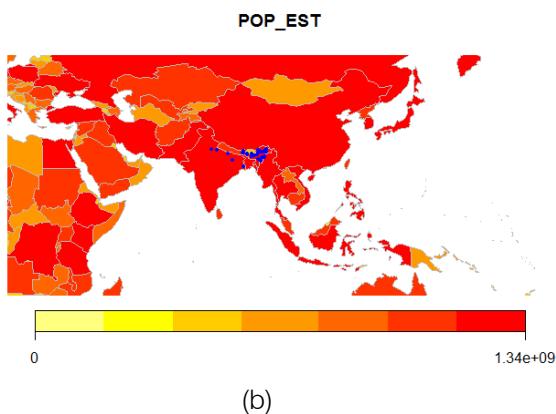
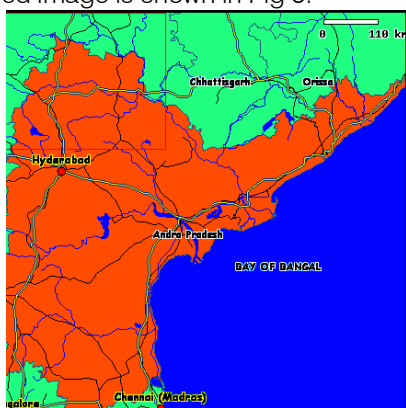
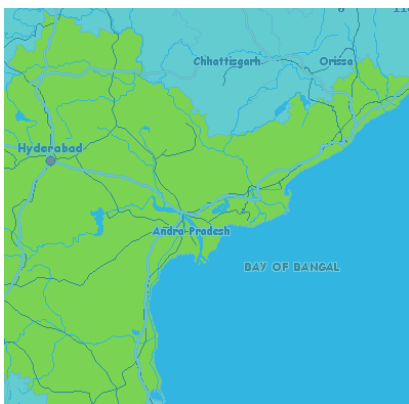


Figure 4: (a) Birds data belonging to Darjeeling area from input dataset in black color (b) Birds data belonging to parts other than Darjeeling area from input dataset marked in blue color

The region of interest for the raster data set and the classified image is shown in Fig 5.



(a)



(b)

Figure 5: (a) Region of Interest from the input raster data set. (b) Classified image with Andhra Pradesh land represented with green and water represented with light blue color.

In this paper the parameters used to evaluate the classification is Accuracy and kappa statistics. The formulae for accuracy, specificity, sensitivity and kappa

statistics are provided by equations (10), (11), (12) and (13)[30][31][32]:

$$\text{Accuracy} = \frac{TP + TN}{(TP + FN + FP + TN)} \times 100 \quad (10)$$

$$\text{Specificity} = \frac{TN}{(TN + FP)} \times 100 \quad (11)$$

$$\text{Sensitivity} = \frac{TP}{(TP + FN)} \times 100 \quad (12)$$

$$\text{Kappa statistics} = \text{Sensitivity} + \text{Specificity} - 1 \quad (13)$$

The efficiency of the proposed SVM classifier is evaluated using the said parameters. The confusion or error matrix view for SVM classifier while classifying the CSV data sets is given in Table 3.

Table 3: Confusion Matrix for CSV data sets

Prediction	Reference	
	Other parts	Darjeeling
Other parts	571	1
Darjeeling	7	116

The confusion matrix or error matrix view for SVM Classifier while classifying raster TIFF data sets is given in Table 4.

Table 4: Confusion Matrix for raster datasets

Prediction	Reference	
	Land	Water
Land	78	0
Water	0	56

Performance Measures using evaluation metrics are specified in Table 5 which are calculated using equations (10), (11), (12) and (13).

Table 5: Performance measures for CSV and raster datasets

Data set type	Accuracy	Kappa Statistics
CSV data sets	98.85	95.97
Raster TIFF data sets	100	100

## V. CONCLUSION

In this paper SVM classification method is used to build a classification model for two datasets. The first data set is of CSV format and the second one is a raster TIFF image. Later the classification model is validated against a test data set which is a subset of the input dataset. The performance of SVM is calculated using kappa statistics and accuracy parameters and it is established that for the given data sets SVM classifies the raster image dataset with better accuracy than the CSV dataset. The SVM classification methodology discussed herein can help in environment monitoring, land use, mineral resource identification, classification of remote sensed data into roads and land etc.. in the future.

## VI. ACKNOWLEDGMENT

We direct our frank appreciativeness to CSV data which was accessed via GBIF data portal,



<http://data.gbif.org/> and image data accessed through <http://maptell.com> for providing us with CSV and raster image data sets. We also thank ANU university for providing all the support in the work conducted.

## REFERENCES RÉFÉRENCES REFERENCIAS

1. E.W.T. Ngai, Yong Hu, Y.H. Wong, Yijun Chen, Xin Sun, The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature, *Decision Support Systems*, Volume 50, Issue 3, February 2011, Pages 559-569, ISSN 01679236, <http://dx.doi.org/10.1016/j.dss.2010.08.006>.
2. Wilson A. Castillo Rojas, Claudio J. Meneses Villegas, Graphical Representation and Exploratory Visualization for Decision Trees in the KDD Process, *Procedia - Social and Behavioral Sciences*, Volume 73, 27 February 2013, Pages 136-144, ISSN 1877-0428, <http://dx.doi.org/10.1016/j.sbspro.2013.02.033>.
3. Borja Calvo, Pedro Larrañaga, José A. Lozano, Learning Bayesian classifiers from positive and unlabeled examples, *Pattern Recognition Letters*, Volume 28, Issue 16, 1 December 2007, Pages 2375-2384, ISSN 01678655, <http://dx.doi.org/10.1016/j.patrec.2007.08.003>.
4. Yugal kumar, G. Sahoo, Analysis of Parametric & Non Parametric Classifiers for Classification Technique using WEKA, *I.J. Information Technology and Computer Science*, 2012, 7, 43-49 Published Online July 2012 in MECS DOI:10.5815/ijitcs.2012.07.06a
5. Ismail Hmeidi, Bilal Hawashin, Eyas El-Qawasmeh, Performance of KNN and SVM classifiers on full word Arabic articles, *Advanced Engineering Informatics*, Volume 22, Issue 1, January 2008, Pages 106-111, ISSN 14740346, [10.1016/j.aei.2007.12.001](http://dx.doi.org/10.1016/j.aei.2007.12.001).
6. Jing Hu, Daoliang Li, Qingling Duan, Yueqi Han, Guifen Chen, Xiuli Si, Fish species classification by color, texture and multi-class support vector machine using computer vision, *Computers and Electronics in Agriculture*, Volume 88, October 2012, Pages 133-140, ISSN 01681699, [10.1016/j.compag.2012.07.008](http://dx.doi.org/10.1016/j.compag.2012.07.008).
7. Xiaowei Yang, Qiaozhen Yu, Lifang He, Tengjiao Guo, The one-against-all partition based binary tree support vector machine algorithms for multi-class classification, *Neuro computing*, Volume 113, 3 August 2013, Pages 17, ISSN 09252312, <http://dx.doi.org/10.1016/j.neucom.2012.12.048>.
8. Yang Shao, Ross S. Lunetta, Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points, *ISPRS Journal of Photogrammetry and Remote Sensing*, Volume 70, June 2012, Pages 7887, ISSN 0924716, [10.1016/j.isprsjprs.2012.04.001](http://dx.doi.org/10.1016/j.isprsjprs.2012.04.001).
9. Rajasekhar, N.; Babu, S.J.; Rajinikanth, T.V., "Magnetic resonance brain images classification using linear kernel based Support Vector Machine, "Engineering (NUICONE), 2012 Nirma University International Conference on , vol., no pp.1,5, 6-8 Dec 2012 doi:10.1109/NUICONE.2012.6493213.
10. Yang Shao, Ross S. Lunetta, Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points, *ISPRS Journal of Photogrammetry and Remote Sensing*, Volume 70, June 2012, Pages 78-87, ISSN 0924-2716, <http://dx.doi.org/10.1016/j.isprsjprs.2012.04.001>.
11. Hongji Lin, Han Lin, Weibin Chen, Study on Recognition of Bird Species in Minjiang River Estuary Wetland, *Procedia Environmental Sciences*, Volume 10, Part C, 2011, Pages 2478-2483, ISSN 1878-0296, <http://dx.doi.org/10.1016/j.proenv.2011.09.386>.
12. Baoping Tang, Tao Song, Feng Li, Lei Deng, Fault diagnosis for a wind turbine transmission system based on manifold learning and Shannon wavelet support vector machine, *Renewable Energy*, Volume 62, February 2014, Pages 1-9, ISSN 0960-1481, <http://dx.doi.org/10.1016/j.renene.2013.06.025>.
13. A.D. Dileep, C. Chandra Sekhar, Class-specific GMM based intermediate matching kernel for classification of varying length patterns of long duration speech using support vector machines, *Speech Communication*, Volume 57, February 2014, Pages 126-143, ISSN 01676393, <http://dx.doi.org/10.1016/j.specom.2013.09.010>.
14. Hsun-Jung Cho, Ming-Te Tseng, A support vector machine approach to CMOS-based radar signal processing for vehicle classification and speed estimation, *Mathematical and Computer Modeling*, Volume 58, Issues 1-2, July 2013, Pages 438-448, ISSN 08957177, <http://dx.doi.org/10.1016/j.mcm.2012.11.003>.
15. Lam Hong Lee, Rajprasad Rajkumar, Lai Hung Lo, Chin Heng Wan, Dino Isa, Oil and gas pipeline failure prediction system using long range ultrasonic transducers and Euclidean-Support Vector Machines classification approach, *Expert Systems with Applications*, Volume 40, Issue 6, May 2013, Pages 1925-1934, ISSN 0957-4174, <http://dx.doi.org/10.1016/j.eswa.2012.10.006>.
16. Elias Zintzaras, Axel Kowald, Forest classification trees and forest support vector machines algorithms: Demonstration using microarray data, *Computers in Biology and Medicine*, Volume 40, Issue 5, May 2010, Pages 519-524, ISSN 00104825, <http://dx.doi.org/10.1016/j.combiomed.2010.03.006>.
17. Xinjun Peng, Yifei Wang, Dong Xu, Structural twin parametric-margin support vector machine for binary classification, *Knowledge-Based Systems*, Volume 49, September 2013, Pages 637-2, ISSN 09507051, <http://dx.doi.org/10.1016/j.knosys.2013.04.013>.

18. Maysam Abedi, Gholam-Hossain Norouzi, Abbas Bahroudi, Support vector machine for multi-classification of mineral prospectivity areas, *Computers & Geosciences*, Volume 46, September 2012, Pages 2722-283, ISSN 0098004, <http://dx.doi.org/10.1016/j.cageo.2011.12.014>.
19. Liu, Mingjun Wang, Jun Wang, Duo Li, Comparison of random forest, support vector machine and back propagation neural network for electronic tongue data classification: Application to the recognition of orange beverage and Chinese vinegar, *Sensors and Actuators B: Chemical*, Volume 177, February 2013, Pages 970-980, ISSN 09254005, <http://dx.doi.org/10.1016/j.snb.2012.11.071>.
20. F. Löw, U. Michel, S. Dech, C. Conrad, Impact of feature selection on the accuracy and spatial uncertainty of per-field crop classification using Support Vector Machines, *ISPRS Journal of Photogrammetry and Remote Sensing*, Volume 85, November 2013, Pages 1021-119, ISSN 09242716, <http://dx.doi.org/10.1016/j.isprsjprs.2013.08.007>.
21. Steve R. Gunn, Support Vector Machines for Classification and Regression, A Technical Report Faculty of Engineering, Science and Mathematics School of Electronics and Computer Science, University Of South Hampton, May 1998
22. Till Rumpf, Christoph Römer, Martin Weis, Markus Sökefeld, Roland Gerhards, Lutz Plümer, Sequential support vector machine classification for small-grain weed species discrimination with special regard to *Cirsium arvense* and *Galium aparine*, *Computers and Electronics in Agriculture*, Volume 80, January 2012, Pages 89-96, ISSN 01681699, <http://dx.doi.org/10.1016/j.compag.2011.10.018>.
23. Hassiba Nemmour, Youcef Chibani, Multiple support vector machines for land cover change detection: An application for mapping urban extensions, *ISPRS Journal of Photogrammetry and Remote Sensing*, Volume 61, Issue 2, November 2006, Pages 125-133, ISSN 09242716, <http://dx.doi.org/10.1016/j.isprsjprs.2006.09.004>.
24. Sujit Narwade, Mohit Kalra, Rajkumar Jagdish, Divya Varier, Sagar Satpute, Gautam Talukdar (2011), Collaborative work of Environmental Information System (ENVIS) Centre and Important Bird Areas Programmes-Indian Bird Conservation Network (IBA-IBCN) projects of the BNHS, India, Zookeys, Narwade, S., et al (2011): Literature based species occurrence data of birds of Northeast India. [www.maptell.com/](http://www.maptell.com/).
25. D. Lu & Q. Weng (2007): A survey of image classification methods and techniques for improving classification performance, *International Journal of Remote Sensing*, 28:5, 823-870, <http://dx.doi.org/10.1080/01431160600746456>
26. S.N. Jeyanthi, Efficient Classification Algorithms using SVMs for Large Datasets, A Project Report Submitted in partial fulfillment of the requirements for the Degree of Master of Technology in Computational Science, Supercomputer Education and Research Center, IISC, BANGALORE, INDIA, June 2007
27. R Core Team}, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2013, <http://www.R-project.org/www.orefeotoolbox.org/otb/monteverdi.html>
28. Jennifer A. Taylor, Alicia V. Lacovara, Gordon S. Smith, Ravi Pandian, Mark Lehto, Near-miss narratives from the fire service: A Bayesian analysis, *Accident Analysis & Prevention*, Volume 62, January 2014, Pages 119-129, ISSN 00014575, <http://dx.doi.org/10.1016/j.aap.2013.09.012>.
29. David J. Rogers, Jonathan E. Suk, Jan C. Semenza, Using global maps to predict the risk of dengue in Europe, *Acta Tropica*, Volume 129, January 2014, Pages 1-14, ISSN 0001-706X, <http://dx.doi.org/10.1016/j.actatropica.2013.08.008>.
30. Rafael Pino-Mejías, María Dolores Cubiles-de-la-Vega, María Anaya-Romero, Antonio Pascual-Acosta, Antonio Jordán-López, Nicolás Bellinfante-Crocci, Predicting the potential habitat of oaks with data mining models and the R system, *Environmental Modelling & Software*, Volume 25, Issue 7, July 2010, Pages 826-836, ISSN 1364-8152, <http://dx.doi.org/10.1016/j.envsoft.2010.01.004>.

GLOBAL JOURNALS INC. (US) GUIDELINES HANDBOOK 2014

---

[WWW.GLOBALJOURNALS.ORG](http://WWW.GLOBALJOURNALS.ORG)

## FELLOWS

### FELLOW OF ASSOCIATION OF RESEARCH SOCIETY IN COMPUTING (FARSC)

Global Journals Incorporate (USA) is accredited by Open Association of Research Society (OARS), U.S.A and in turn, awards “FARSC” title to individuals. The 'FARSC' title is accorded to a selected professional after the approval of the Editor-in-Chief/Editorial Board Members/Dean.



- The “FARSC” is a dignified title which is accorded to a person’s name viz. Dr. John E. Hall, Ph.D., FARSC or William Walldroff, M.S., FARSC.

FARSC accrediting is an honor. It authenticates your research activities. After recognition as FARSC, you can add 'FARSC' title with your name as you use this recognition as additional suffix to your status. This will definitely enhance and add more value and repute to your name. You may use it on your professional Counseling Materials such as CV, Resume, and Visiting Card etc.

*The following benefits can be availed by you only for next three years from the date of certification:*



FARSC designated members are entitled to avail a 40% discount while publishing their research papers (of a single author) with Global Journals Incorporation (USA), if the same is accepted by Editorial Board/Peer Reviewers. If you are a main author or co-author in case of multiple authors, you will be entitled to avail discount of 10%.

Once FARSC title is accorded, the Fellow is authorized to organize a symposium/seminar/conference on behalf of Global Journal Incorporation (USA). The Fellow can also participate in conference/seminar/symposium organized by another institution as representative of Global Journal. In both the cases, it is mandatory for him to discuss with us and obtain our consent.



You may join as member of the Editorial Board of Global Journals Incorporation (USA) after successful completion of three years as Fellow and as Peer Reviewer. In addition, it is also desirable that you should organize seminar/symposium/conference at least once.

We shall provide you intimation regarding launching of e-version of journal of your stream time to time. This may be utilized in your library for the enrichment of knowledge of your students as well as it can also be helpful for the concerned faculty members.

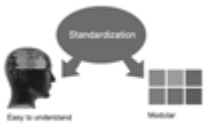




Journals Research  
inducing researches

The FARSC can go through standards of OARS. You can also play vital role if you have any suggestions so that proper amendment can take place to improve the same for the benefit of entire research community.

As FARSC, you will be given a renowned, secure and free professional email address with 100 GB of space e.g. [johnhall@globaljournals.org](mailto:johnhall@globaljournals.org). This will include Webmail, Spam Assassin, Email Forwarders, Auto-Responders, Email Delivery Route tracing, etc.



The FARSC will be eligible for a free application of standardization of their researches. Standardization of research will be subject to acceptability within stipulated norms as the next step after publishing in a journal. We shall depute a team of specialized research professionals who will render their services for elevating your researches to next higher level, which is worldwide open standardization.

The FARSC member can apply for grading and certification of standards of their educational and Institutional Degrees to Open Association of Research, Society U.S.A. Once you are designated as FARSC, you may send us a scanned copy of all of your credentials. OARS will verify, grade and certify them. This will be based on your academic records, quality of research papers published by you, and some more criteria. After certification of all your credentials by OARS, they will be published on your Fellow Profile link on website <https://associationofresearch.org> which will be helpful to upgrade the dignity.



The FARSC members can avail the benefits of free research podcasting in Global Research Radio with their research documents. After publishing the work, (including published elsewhere worldwide with proper authorization) you can upload your research paper with your recorded voice or you can utilize chargeable services of our professional RJs to record your paper in their voice on request.

The FARSC member also entitled to get the benefits of free research podcasting of their research documents through video clips. We can also streamline your conference videos and display your slides/ online slides and online research video clips at reasonable charges, on request.





The FARSC is eligible to earn from sales proceeds of his/her researches/reference/review Books or literature, while publishing with Global Journals. The FARSC can decide whether he/she would like to publish his/her research in a closed manner. In this case, whenever readers purchase that individual research paper for reading, maximum 60% of its profit earned as royalty by Global Journals, will be credited to his/her bank account. The entire entitled amount will be credited to his/her bank account exceeding limit of minimum fixed balance. There is no minimum time limit for collection. The FARSC member can decide its price and we can help in making the right decision.

The FARSC member is eligible to join as a paid peer reviewer at Global Journals Incorporation (USA) and can get remuneration of 15% of author fees, taken from the author of a respective paper. After reviewing 5 or more papers you can request to transfer the amount to your bank account.



## MEMBER OF ASSOCIATION OF RESEARCH SOCIETY IN COMPUTING (MARSC)

The ' MARSC ' title is accorded to a selected professional after the approval of the Editor-in-Chief / Editorial Board Members/Dean.

The "MARSC" is a dignified ornament which is accorded to a person's name viz. Dr. John E. Hall, Ph.D., MARSC or William Walldroff, M.S., MARSC.



MARSC accrediting is an honor. It authenticates your research activities. After becoming MARSC, you can add 'MARSC' title with your name as you use this recognition as additional suffix to your status. This will definitely enhance and add more value and repute to your name. You may use it on your professional Counseling Materials such as CV, Resume, Visiting Card and Name Plate etc.

*The following benefits can be availed by you only for next three years from the date of certification.*



MARSC designated members are entitled to avail a 25% discount while publishing their research papers (of a single author) in Global Journals Inc., if the same is accepted by our Editorial Board and Peer Reviewers. If you are a main author or co-author of a group of authors, you will get discount of 10%.

As MARSC, you will be given a renowned, secure and free professional email address with 30 GB of space e.g. [johnhall@globaljournals.org](mailto:johnhall@globaljournals.org). This will include Webmail, Spam Assassin, Email Forwarders, Auto-Responders, Email Delivery Route tracing, etc.





We shall provide you intimation regarding launching of e-version of journal of your stream time to time. This may be utilized in your library for the enrichment of knowledge of your students as well as it can also be helpful for the concerned faculty members.

The MARSC member can apply for approval, grading and certification of standards of their educational and Institutional Degrees to Open Association of Research, Society U.S.A.



Once you are designated as MARSC, you may send us a scanned copy of all of your credentials. OARS will verify, grade and certify them. This will be based on your academic records, quality of research papers published by you, and some more criteria.

It is mandatory to read all terms and conditions carefully.



## AUXILIARY MEMBERSHIPS

### Institutional Fellow of Open Association of Research Society (USA)-OARS (USA)

Global Journals Incorporation (USA) is accredited by Open Association of Research Society, U.S.A (OARS) and in turn, affiliates research institutions as “Institutional Fellow of Open Association of Research Society” (IFOARS).



The “FARSC” is a dignified title which is accorded to a person’s name viz. Dr. John E. Hall, Ph.D., FARSC or William Walldroff, M.S., FARSC.

The IFOARS institution is entitled to form a Board comprised of one Chairperson and three to five board members preferably from different streams. The Board will be recognized as “Institutional Board of Open Association of Research Society”-(IBOARS).

*The Institute will be entitled to following benefits:*



The IBOARS can initially review research papers of their institute and recommend them to publish with respective journal of Global Journals. It can also review the papers of other institutions after obtaining our consent. The second review will be done by peer reviewer of Global Journals Incorporation (USA) The Board is at liberty to appoint a peer reviewer with the approval of chairperson after consulting us.

The author fees of such paper may be waived off up to 40%.

The Global Journals Incorporation (USA) at its discretion can also refer double blind peer reviewed paper at their end to the board for the verification and to get recommendation for final stage of acceptance of publication.



The IBOARS can organize symposium/seminar/conference in their country on behalf of Global Journals Incorporation (USA)-OARS (USA). The terms and conditions can be discussed separately.

The Board can also play vital role by exploring and giving valuable suggestions regarding the Standards of “Open Association of Research Society, U.S.A (OARS)” so that proper amendment can take place for the benefit of entire research community. We shall provide details of particular standard only on receipt of request from the Board.



Journals Research  
inducing researches

The board members can also join us as Individual Fellow with 40% discount on total fees applicable to Individual Fellow. They will be entitled to avail all the benefits as declared. Please visit Individual Fellow-sub menu of GlobalJournals.org to have more relevant details.



We shall provide you intimation regarding launching of e-version of journal of your stream time to time. This may be utilized in your library for the enrichment of knowledge of your students as well as it can also be helpful for the concerned faculty members.



After nomination of your institution as “Institutional Fellow” and constantly functioning successfully for one year, we can consider giving recognition to your institute to function as Regional/Zonal office on our behalf.

The board can also take up the additional allied activities for betterment after our consultation.

### **The following entitlements are applicable to individual Fellows:**

Open Association of Research Society, U.S.A (OARS) By-laws states that an individual Fellow may use the designations as applicable, or the corresponding initials. The Credentials of individual Fellow and Associate designations signify that the individual has gained knowledge of the fundamental concepts. One is magnanimous and proficient in an expertise course covering the professional code of conduct, and follows recognized standards of practice.



Open Association of Research Society (US)/ Global Journals Incorporation (USA), as described in Corporate Statements, are educational, research publishing and professional membership organizations. Achieving our individual Fellow or Associate status is based mainly on meeting stated educational research requirements.

Disbursement of 40% Royalty earned through Global Journals : Researcher = 50%, Peer Reviewer = 37.50%, Institution = 12.50% E.g. Out of 40%, the 20% benefit should be passed on to researcher, 15 % benefit towards remuneration should be given to a reviewer and remaining 5% is to be retained by the institution.



We shall provide print version of 12 issues of any three journals [as per your requirement] out of our 38 journals worth \$ 2376 USD.

### **Other:**

**The individual Fellow and Associate designations accredited by Open Association of Research Society (US) credentials signify guarantees following achievements:**

- The professional accredited with Fellow honor, is entitled to various benefits viz. name, fame, honor, regular flow of income, secured bright future, social status etc.



- In addition to above, if one is single author, then entitled to 40% discount on publishing research paper and can get 10% discount if one is co-author or main author among group of authors.
- The Fellow can organize symposium/seminar/conference on behalf of Global Journals Incorporation (USA) and he/she can also attend the same organized by other institutes on behalf of Global Journals.
- The Fellow can become member of Editorial Board Member after completing 3yrs.
- The Fellow can earn 60% of sales proceeds from the sale of reference/review books/literature/publishing of research paper.
- Fellow can also join as paid peer reviewer and earn 15% remuneration of author charges and can also get an opportunity to join as member of the Editorial Board of Global Journals Incorporation (USA)
- • This individual has learned the basic methods of applying those concepts and techniques to common challenging situations. This individual has further demonstrated an in-depth understanding of the application of suitable techniques to a particular area of research practice.

**Note :**

“

- In future, if the board feels the necessity to change any board member, the same can be done with the consent of the chairperson along with anyone board member without our approval.
- In case, the chairperson needs to be replaced then consent of 2/3rd board members are required and they are also required to jointly pass the resolution copy of which should be sent to us. In such case, it will be compulsory to obtain our approval before replacement.
- In case of “Difference of Opinion [if any]” among the Board members, our decision will be final and binding to everyone.

”



## PROCESS OF SUBMISSION OF RESEARCH PAPER

---

The Area or field of specialization may or may not be of any category as mentioned in 'Scope of Journal' menu of the GlobalJournals.org website. There are 37 Research Journal categorized with Six parental Journals GJCST, GJMR, GJRE, GJMBR, GJSFR, GJHSS. For Authors should prefer the mentioned categories. There are three widely used systems UDC, DDC and LCC. The details are available as 'Knowledge Abstract' at Home page. The major advantage of this coding is that, the research work will be exposed to and shared with all over the world as we are being abstracted and indexed worldwide.

The paper should be in proper format. The format can be downloaded from first page of 'Author Guideline' Menu. The Author is expected to follow the general rules as mentioned in this menu. The paper should be written in MS-Word Format (\*.DOC,\*.DOCX).

The Author can submit the paper either online or offline. The authors should prefer online submission.Online Submission: There are three ways to submit your paper:

**(A) (I) First, register yourself using top right corner of Home page then Login. If you are already registered, then login using your username and password.**

**(II) Choose corresponding Journal.**

**(III) Click 'Submit Manuscript'. Fill required information and Upload the paper.**

**(B) If you are using Internet Explorer, then Direct Submission through Homepage is also available.**

**(C) If these two are not convenient, and then email the paper directly to dean@globaljournals.org.**

Offline Submission: Author can send the typed form of paper by Post. However, online submission should be preferred.



# PREFERRED AUTHOR GUIDELINES

## MANUSCRIPT STYLE INSTRUCTION (Must be strictly followed)

Page Size: 8.27" X 11"

- Left Margin: 0.65
- Right Margin: 0.65
- Top Margin: 0.75
- Bottom Margin: 0.75
- Font type of all text should be Swis 721 Lt BT.
- Paper Title should be of Font Size 24 with one Column section.
- Author Name in Font Size of 11 with one column as of Title.
- Abstract Font size of 9 Bold, "Abstract" word in Italic Bold.
- Main Text: Font size 10 with justified two columns section
- Two Column with Equal Column with of 3.38 and Gaping of .2
- First Character must be three lines Drop capped.
- Paragraph before Spacing of 1 pt and After of 0 pt.
- Line Spacing of 1 pt
- Large Images must be in One Column
- Numbering of First Main Headings (Heading 1) must be in Roman Letters, Capital Letter, and Font Size of 10.
- Numbering of Second Main Headings (Heading 2) must be in Alphabets, Italic, and Font Size of 10.

**You can use your own standard format also.**

### Author Guidelines:

1. General,
2. Ethical Guidelines,
3. Submission of Manuscripts,
4. Manuscript's Category,
5. Structure and Format of Manuscript,
6. After Acceptance.

### 1. GENERAL

Before submitting your research paper, one is advised to go through the details as mentioned in following heads. It will be beneficial, while peer reviewer justify your paper for publication.

### Scope

The Global Journals Inc. (US) welcome the submission of original paper, review paper, survey article relevant to the all the streams of Philosophy and knowledge. The Global Journals Inc. (US) is parental platform for Global Journal of Computer Science and Technology, Researches in Engineering, Medical Research, Science Frontier Research, Human Social Science, Management, and Business organization. The choice of specific field can be done otherwise as following in Abstracting and Indexing Page on this Website. As the all Global

Journals Inc. (US) are being abstracted and indexed (in process) by most of the reputed organizations. Topics of only narrow interest will not be accepted unless they have wider potential or consequences.

## 2. ETHICAL GUIDELINES

Authors should follow the ethical guidelines as mentioned below for publication of research paper and research activities.

Papers are accepted on strict understanding that the material in whole or in part has not been, nor is being, considered for publication elsewhere. If the paper once accepted by Global Journals Inc. (US) and Editorial Board, will become the copyright of the Global Journals Inc. (US).

**Authorship: The authors and coauthors should have active contribution to conception design, analysis and interpretation of findings. They should critically review the contents and drafting of the paper. All should approve the final version of the paper before submission**

The Global Journals Inc. (US) follows the definition of authorship set up by the Global Academy of Research and Development. According to the Global Academy of R&D authorship, criteria must be based on:

- 1) Substantial contributions to conception and acquisition of data, analysis and interpretation of the findings.
- 2) Drafting the paper and revising it critically regarding important academic content.
- 3) Final approval of the version of the paper to be published.

All authors should have been credited according to their appropriate contribution in research activity and preparing paper. Contributors who do not match the criteria as authors may be mentioned under Acknowledgement.

Acknowledgements: Contributors to the research other than authors credited should be mentioned under acknowledgement. The specifications of the source of funding for the research if appropriate can be included. Suppliers of resources may be mentioned along with address.

**Appeal of Decision: The Editorial Board's decision on publication of the paper is final and cannot be appealed elsewhere.**

**Permissions: It is the author's responsibility to have prior permission if all or parts of earlier published illustrations are used in this paper.**

Please mention proper reference and appropriate acknowledgements wherever expected.

If all or parts of previously published illustrations are used, permission must be taken from the copyright holder concerned. It is the author's responsibility to take these in writing.

Approval for reproduction/modification of any information (including figures and tables) published elsewhere must be obtained by the authors/copyright holders before submission of the manuscript. Contributors (Authors) are responsible for any copyright fee involved.

## 3. SUBMISSION OF MANUSCRIPTS

Manuscripts should be uploaded via this online submission page. The online submission is most efficient method for submission of papers, as it enables rapid distribution of manuscripts and consequently speeds up the review procedure. It also enables authors to know the status of their own manuscripts by emailing us. Complete instructions for submitting a paper is available below.

Manuscript submission is a systematic procedure and little preparation is required beyond having all parts of your manuscript in a given format and a computer with an Internet connection and a Web browser. Full help and instructions are provided on-screen. As an author, you will be prompted for login and manuscript details as Field of Paper and then to upload your manuscript file(s) according to the instructions.



To avoid postal delays, all transaction is preferred by e-mail. A finished manuscript submission is confirmed by e-mail immediately and your paper enters the editorial process with no postal delays. When a conclusion is made about the publication of your paper by our Editorial Board, revisions can be submitted online with the same procedure, with an occasion to view and respond to all comments.

Complete support for both authors and co-author is provided.

#### 4. MANUSCRIPT'S CATEGORY

Based on potential and nature, the manuscript can be categorized under the following heads:

Original research paper: Such papers are reports of high-level significant original research work.

Review papers: These are concise, significant but helpful and decisive topics for young researchers.

Research articles: These are handled with small investigation and applications.

Research letters: The letters are small and concise comments on previously published matters.

#### 5. STRUCTURE AND FORMAT OF MANUSCRIPT

The recommended size of original research paper is less than seven thousand words, review papers fewer than seven thousands words also. Preparation of research paper or how to write research paper, are major hurdle, while writing manuscript. The research articles and research letters should be fewer than three thousand words, the structure original research paper; sometime review paper should be as follows:

**Papers:** These are reports of significant research (typically less than 7000 words equivalent, including tables, figures, references), and comprise:

- (a) Title should be relevant and commensurate with the theme of the paper.
- (b) A brief Summary, "Abstract" (less than 150 words) containing the major results and conclusions.
- (c) Up to ten keywords, that precisely identifies the paper's subject, purpose, and focus.
- (d) An Introduction, giving necessary background excluding subheadings; objectives must be clearly declared.
- (e) Resources and techniques with sufficient complete experimental details (wherever possible by reference) to permit repetition; sources of information must be given and numerical methods must be specified by reference, unless non-standard.
- (f) Results should be presented concisely, by well-designed tables and/or figures; the same data may not be used in both; suitable statistical data should be given. All data must be obtained with attention to numerical detail in the planning stage. As reproduced design has been recognized to be important to experiments for a considerable time, the Editor has decided that any paper that appears not to have adequate numerical treatments of the data will be returned un-refereed;
- (g) Discussion should cover the implications and consequences, not just recapitulating the results; conclusions should be summarizing.
- (h) Brief Acknowledgements.
- (i) References in the proper form.

Authors should very cautiously consider the preparation of papers to ensure that they communicate efficiently. Papers are much more likely to be accepted, if they are cautiously designed and laid out, contain few or no errors, are summarizing, and be conventional to the approach and instructions. They will in addition, be published with much less delays than those that require much technical and editorial correction.



The Editorial Board reserves the right to make literary corrections and to make suggestions to improve brevity.

It is vital, that authors take care in submitting a manuscript that is written in simple language and adheres to published guidelines.

## Format

*Language: The language of publication is UK English. Authors, for whom English is a second language, must have their manuscript efficiently edited by an English-speaking person before submission to make sure that, the English is of high excellence. It is preferable, that manuscripts should be professionally edited.*

Standard Usage, Abbreviations, and Units: Spelling and hyphenation should be conventional to The Concise Oxford English Dictionary. Statistics and measurements should at all times be given in figures, e.g. 16 min, except for when the number begins a sentence. When the number does not refer to a unit of measurement it should be spelt in full unless, it is 160 or greater.

Abbreviations supposed to be used carefully. The abbreviated name or expression is supposed to be cited in full at first usage, followed by the conventional abbreviation in parentheses.

Metric SI units are supposed to generally be used excluding where they conflict with current practice or are confusing. For illustration, 1.4 l rather than  $1.4 \times 10^{-3} \text{ m}^3$ , or 4 mm somewhat than  $4 \times 10^{-3} \text{ m}$ . Chemical formula and solutions must identify the form used, e.g. anhydrous or hydrated, and the concentration must be in clearly defined units. Common species names should be followed by underlines at the first mention. For following use the generic name should be constricted to a single letter, if it is clear.

## Structure

All manuscripts submitted to Global Journals Inc. (US), ought to include:

Title: The title page must carry an instructive title that reflects the content, a running title (less than 45 characters together with spaces), names of the authors and co-authors, and the place(s) wherever the work was carried out. The full postal address in addition with the e-mail address of related author must be given. Up to eleven keywords or very brief phrases have to be given to help data retrieval, mining and indexing.

*Abstract, used in Original Papers and Reviews:*

### Optimizing Abstract for Search Engines

Many researchers searching for information online will use search engines such as Google, Yahoo or similar. By optimizing your paper for search engines, you will amplify the chance of someone finding it. This in turn will make it more likely to be viewed and/or cited in a further work. Global Journals Inc. (US) have compiled these guidelines to facilitate you to maximize the web-friendliness of the most public part of your paper.

### Key Words

A major linchpin in research work for the writing research paper is the keyword search, which one will employ to find both library and Internet resources.

One must be persistent and creative in using keywords. An effective keyword search requires a strategy and planning a list of possible keywords and phrases to try.

Search engines for most searches, use Boolean searching, which is somewhat different from Internet searches. The Boolean search uses "operators," words (and, or, not, and near) that enable you to expand or narrow your affords. Tips for research paper while preparing research paper are very helpful guideline of research paper.

Choice of key words is first tool of tips to write research paper. Research paper writing is an art. A few tips for deciding as strategically as possible about keyword search:



- One should start brainstorming lists of possible keywords before even begin searching. Think about the most important concepts related to research work. Ask, "What words would a source have to include to be truly valuable in research paper?" Then consider synonyms for the important words.
- It may take the discovery of only one relevant paper to let steer in the right keyword direction because in most databases, the keywords under which a research paper is abstracted are listed with the paper.
- One should avoid outdated words.

Keywords are the key that opens a door to research work sources. Keyword searching is an art in which researcher's skills are bound to improve with experience and time.

Numerical Methods: Numerical methods used should be clear and, where appropriate, supported by references.

*Acknowledgements: Please make these as concise as possible.*

#### References

References follow the Harvard scheme of referencing. References in the text should cite the authors' names followed by the time of their publication, unless there are three or more authors when simply the first author's name is quoted followed by et al. unpublished work has to only be cited where necessary, and only in the text. Copies of references in press in other journals have to be supplied with submitted typescripts. It is necessary that all citations and references be carefully checked before submission, as mistakes or omissions will cause delays.

References to information on the World Wide Web can be given, but only if the information is available without charge to readers on an official site. Wikipedia and Similar websites are not allowed where anyone can change the information. Authors will be asked to make available electronic copies of the cited information for inclusion on the Global Journals Inc. (US) homepage at the judgment of the Editorial Board.

The Editorial Board and Global Journals Inc. (US) recommend that, citation of online-published papers and other material should be done via a DOI (digital object identifier). If an author cites anything, which does not have a DOI, they run the risk of the cited material not being noticeable.

The Editorial Board and Global Journals Inc. (US) recommend the use of a tool such as Reference Manager for reference management and formatting.

#### Tables, Figures and Figure Legends

*Tables: Tables should be few in number, cautiously designed, uncrowned, and include only essential data. Each must have an Arabic number, e.g. Table 4, a self-explanatory caption and be on a separate sheet. Vertical lines should not be used.*

*Figures: Figures are supposed to be submitted as separate files. Always take in a citation in the text for each figure using Arabic numbers, e.g. Fig. 4. Artwork must be submitted online in electronic form by e-mailing them.*

#### Preparation of Electronic Figures for Publication

Even though low quality images are sufficient for review purposes, print publication requires high quality images to prevent the final product being blurred or fuzzy. Submit (or e-mail) EPS (line art) or TIFF (halftone/photographs) files only. MS PowerPoint and Word Graphics are unsuitable for printed pictures. Do not use pixel-oriented software. Scans (TIFF only) should have a resolution of at least 350 dpi (halftone) or 700 to 1100 dpi (line drawings) in relation to the imitation size. Please give the data for figures in black and white or submit a Color Work Agreement Form. EPS files must be saved with fonts embedded (and with a TIFF preview, if possible).

For scanned images, the scanning resolution (at final image size) ought to be as follows to ensure good reproduction: line art: >650 dpi; halftones (including gel photographs) : >350 dpi; figures containing both halftone and line images: >650 dpi.

Color Charges: It is the rule of the Global Journals Inc. (US) for authors to pay the full cost for the reproduction of their color artwork. Hence, please note that, if there is color artwork in your manuscript when it is accepted for publication, we would require you to complete and return a color work agreement form before your paper can be published.





*Figure Legends: Self-explanatory legends of all figures should be incorporated separately under the heading 'Legends to Figures'. In the full-text online edition of the journal, figure legends may possibly be truncated in abbreviated links to the full screen version. Therefore, the first 100 characters of any legend should notify the reader, about the key aspects of the figure.*

## **6. AFTER ACCEPTANCE**

Upon approval of a paper for publication, the manuscript will be forwarded to the dean, who is responsible for the publication of the Global Journals Inc. (US).

### **6.1 Proof Corrections**

The corresponding author will receive an e-mail alert containing a link to a website or will be attached. A working e-mail address must therefore be provided for the related author.

Acrobat Reader will be required in order to read this file. This software can be downloaded

(Free of charge) from the following website:

[www.adobe.com/products/acrobat/readstep2.html](http://www.adobe.com/products/acrobat/readstep2.html). This will facilitate the file to be opened, read on screen, and printed out in order for any corrections to be added. Further instructions will be sent with the proof.

Proofs must be returned to the dean at [dean@globaljournals.org](mailto:dean@globaljournals.org) within three days of receipt.

As changes to proofs are costly, we inquire that you only correct typesetting errors. All illustrations are retained by the publisher. Please note that the authors are responsible for all statements made in their work, including changes made by the copy editor.

### **6.2 Early View of Global Journals Inc. (US) (Publication Prior to Print)**

The Global Journals Inc. (US) are enclosed by our publishing's Early View service. Early View articles are complete full-text articles sent in advance of their publication. Early View articles are absolute and final. They have been completely reviewed, revised and edited for publication, and the authors' final corrections have been incorporated. Because they are in final form, no changes can be made after sending them. The nature of Early View articles means that they do not yet have volume, issue or page numbers, so Early View articles cannot be cited in the conventional way.

### **6.3 Author Services**

Online production tracking is available for your article through Author Services. Author Services enables authors to track their article - once it has been accepted - through the production process to publication online and in print. Authors can check the status of their articles online and choose to receive automated e-mails at key stages of production. The authors will receive an e-mail with a unique link that enables them to register and have their article automatically added to the system. Please ensure that a complete e-mail address is provided when submitting the manuscript.

### **6.4 Author Material Archive Policy**

Please note that if not specifically requested, publisher will dispose off hardcopy & electronic information submitted, after the two months of publication. If you require the return of any information submitted, please inform the Editorial Board or dean as soon as possible.

### **6.5 Offprint and Extra Copies**

A PDF offprint of the online-published article will be provided free of charge to the related author, and may be distributed according to the Publisher's terms and conditions. Additional paper offprint may be ordered by emailing us at: [editor@globaljournals.org](mailto:editor@globaljournals.org).

You must strictly follow above Author Guidelines before submitting your paper or else we will not at all be responsible for any corrections in future in any of the way.



Before start writing a good quality Computer Science Research Paper, let us first understand what is Computer Science Research Paper? So, Computer Science Research Paper is the paper which is written by professionals or scientists who are associated to Computer Science and Information Technology, or doing research study in these areas. If you are novel to this field then you can consult about this field from your supervisor or guide.

#### TECHNIQUES FOR WRITING A GOOD QUALITY RESEARCH PAPER:

**1. Choosing the topic:** In most cases, the topic is searched by the interest of author but it can be also suggested by the guides. You can have several topics and then you can judge that in which topic or subject you are finding yourself most comfortable. This can be done by asking several questions to yourself, like Will I be able to carry our search in this area? Will I find all necessary recourses to accomplish the search? Will I be able to find all information in this field area? If the answer of these types of questions will be "Yes" then you can choose that topic. In most of the cases, you may have to conduct the surveys and have to visit several places because this field is related to Computer Science and Information Technology. Also, you may have to do a lot of work to find all rise and falls regarding the various data of that subject. Sometimes, detailed information plays a vital role, instead of short information.

**2. Evaluators are human:** First thing to remember that evaluators are also human being. They are not only meant for rejecting a paper. They are here to evaluate your paper. So, present your Best.

**3. Think Like Evaluators:** If you are in a confusion or getting demotivated that your paper will be accepted by evaluators or not, then think and try to evaluate your paper like an Evaluator. Try to understand that what an evaluator wants in your research paper and automatically you will have your answer.

**4. Make blueprints of paper:** The outline is the plan or framework that will help you to arrange your thoughts. It will make your paper logical. But remember that all points of your outline must be related to the topic you have chosen.

**5. Ask your Guides:** If you are having any difficulty in your research, then do not hesitate to share your difficulty to your guide (if you have any). They will surely help you out and resolve your doubts. If you can't clarify what exactly you require for your work then ask the supervisor to help you with the alternative. He might also provide you the list of essential readings.

**6. Use of computer is recommended:** As you are doing research in the field of Computer Science, then this point is quite obvious.

**7. Use right software:** Always use good quality software packages. If you are not capable to judge good software then you can lose quality of your paper unknowingly. There are various software programs available to help you, which you can get through Internet.

**8. Use the Internet for help:** An excellent start for your paper can be by using the Google. It is an excellent search engine, where you can have your doubts resolved. You may also read some answers for the frequent question how to write my research paper or find model research paper. From the internet library you can download books. If you have all required books make important reading selecting and analyzing the specified information. Then put together research paper sketch out.

**9. Use and get big pictures:** Always use encyclopedias, Wikipedia to get pictures so that you can go into the depth.

**10. Bookmarks are useful:** When you read any book or magazine, you generally use bookmarks, right! It is a good habit, which helps to not to lose your continuity. You should always use bookmarks while searching on Internet also, which will make your search easier.

**11. Revise what you wrote:** When you write anything, always read it, summarize it and then finalize it.



**12. Make all efforts:** Make all efforts to mention what you are going to write in your paper. That means always have a good start. Try to mention everything in introduction, that what is the need of a particular research paper. Polish your work by good skill of writing and always give an evaluator, what he wants.

**13. Have backups:** When you are going to do any important thing like making research paper, you should always have backup copies of it either in your computer or in paper. This will help you to not to lose any of your important.

**14. Produce good diagrams of your own:** Always try to include good charts or diagrams in your paper to improve quality. Using several and unnecessary diagrams will degrade the quality of your paper by creating "hotchpotch." So always, try to make and include those diagrams, which are made by your own to improve readability and understandability of your paper.

**15. Use of direct quotes:** When you do research relevant to literature, history or current affairs then use of quotes become essential but if study is relevant to science then use of quotes is not preferable.

**16. Use proper verb tense:** Use proper verb tenses in your paper. Use past tense, to present those events that happened. Use present tense to indicate events that are going on. Use future tense to indicate future happening events. Use of improper and wrong tenses will confuse the evaluator. Avoid the sentences that are incomplete.

**17. Never use online paper:** If you are getting any paper on Internet, then never use it as your research paper because it might be possible that evaluator has already seen it or maybe it is outdated version.

**18. Pick a good study spot:** To do your research studies always try to pick a spot, which is quiet. Every spot is not for studies. Spot that suits you choose it and proceed further.

**19. Know what you know:** Always try to know, what you know by making objectives. Else, you will be confused and cannot achieve your target.

**20. Use good quality grammar:** Always use a good quality grammar and use words that will throw positive impact on evaluator. Use of good quality grammar does not mean to use tough words, that for each word the evaluator has to go through dictionary. Do not start sentence with a conjunction. Do not fragment sentences. Eliminate one-word sentences. Ignore passive voice. Do not ever use a big word when a diminutive one would suffice. Verbs have to be in agreement with their subjects. Prepositions are not expressions to finish sentences with. It is incorrect to ever divide an infinitive. Avoid clichés like the disease. Also, always shun irritating alliteration. Use language that is simple and straight forward. put together a neat summary.

**21. Arrangement of information:** Each section of the main body should start with an opening sentence and there should be a changeover at the end of the section. Give only valid and powerful arguments to your topic. You may also maintain your arguments with records.

**22. Never start in last minute:** Always start at right time and give enough time to research work. Leaving everything to the last minute will degrade your paper and spoil your work.

**23. Multitasking in research is not good:** Doing several things at the same time proves bad habit in case of research activity. Research is an area, where everything has a particular time slot. Divide your research work in parts and do particular part in particular time slot.

**24. Never copy others' work:** Never copy others' work and give it your name because if evaluator has seen it anywhere you will be in trouble.

**25. Take proper rest and food:** No matter how many hours you spend for your research activity, if you are not taking care of your health then all your efforts will be in vain. For a quality research, study is must, and this can be done by taking proper rest and food.

**26. Go for seminars:** Attend seminars if the topic is relevant to your research area. Utilize all your resources.



**27. Refresh your mind after intervals:** Try to give rest to your mind by listening to soft music or by sleeping in intervals. This will also improve your memory.

**28. Make colleagues:** Always try to make colleagues. No matter how sharper or intelligent you are, if you make colleagues you can have several ideas, which will be helpful for your research.

**29. Think technically:** Always think technically. If anything happens, then search its reasons, its benefits, and demerits.

**30. Think and then print:** When you will go to print your paper, notice that tables are not be split, headings are not detached from their descriptions, and page sequence is maintained.

**31. Adding unnecessary information:** Do not add unnecessary information, like, I have used MS Excel to draw graph. Do not add irrelevant and inappropriate material. These all will create superfluous. Foreign terminology and phrases are not apropos. One should NEVER take a broad view. Analogy in script is like feathers on a snake. Not at all use a large word when a very small one would be sufficient. Use words properly, regardless of how others use them. Remove quotations. Puns are for kids, not grunt readers. Amplification is a billion times of inferior quality than sarcasm.

**32. Never oversimplify everything:** To add material in your research paper, never go for oversimplification. This will definitely irritate the evaluator. Be more or less specific. Also too, by no means, ever use rhythmic redundancies. Contractions aren't essential and shouldn't be there used. Comparisons are as terrible as clichés. Give up ampersands and abbreviations, and so on. Remove commas, that are, not necessary. Parenthetical words however should be together with this in commas. Understatement is all the time the complete best way to put onward earth-shaking thoughts. Give a detailed literary review.

**33. Report concluded results:** Use concluded results. From raw data, filter the results and then conclude your studies based on measurements and observations taken. Significant figures and appropriate number of decimal places should be used. Parenthetical remarks are prohibitive. Proofread carefully at final stage. In the end give outline to your arguments. Spot out perspectives of further study of this subject. Justify your conclusion by at the bottom of them with sufficient justifications and examples.

**34. After conclusion:** Once you have concluded your research, the next most important step is to present your findings. Presentation is extremely important as it is the definite medium through which your research is going to be in print to the rest of the crowd. Care should be taken to categorize your thoughts well and present them in a logical and neat manner. A good quality research paper format is essential because it serves to highlight your research paper and bring to light all necessary aspects in your research.

## INFORMAL GUIDELINES OF RESEARCH PAPER WRITING

### Key points to remember:

- Submit all work in its final form.
- Write your paper in the form, which is presented in the guidelines using the template.
- Please note the criterion for grading the final paper by peer-reviewers.

### Final Points:

A purpose of organizing a research paper is to let people to interpret your effort selectively. The journal requires the following sections, submitted in the order listed, each section to start on a new page.

The introduction will be compiled from reference matter and will reflect the design processes or outline of basis that direct you to make study. As you will carry out the process of study, the method and process section will be constructed as like that. The result segment will show related statistics in nearly sequential order and will direct the reviewers next to the similar intellectual paths throughout the data that you took to carry out your study. The discussion section will provide understanding of the data and projections as to the implication of the results. The use of good quality references all through the paper will give the effort trustworthiness by representing an alertness of prior workings.



Writing a research paper is not an easy job no matter how trouble-free the actual research or concept. Practice, excellent preparation, and controlled record keeping are the only means to make straightforward the progression.

### **General style:**

Specific editorial column necessities for compliance of a manuscript will always take over from directions in these general guidelines.

To make a paper clear

- Adhere to recommended page limits

Mistakes to evade

- Insertion a title at the foot of a page with the subsequent text on the next page
- Separating a table/chart or figure - impound each figure/table to a single page
- Submitting a manuscript with pages out of sequence

In every sections of your document

- Use standard writing style including articles ("a", "the," etc.)
- Keep on paying attention on the research topic of the paper
- Use paragraphs to split each significant point (excluding for the abstract)
- Align the primary line of each section
- Present your points in sound order
- Use present tense to report well accepted
- Use past tense to describe specific results
- Shun familiar wording, don't address the reviewer directly, and don't use slang, slang language, or superlatives
- Shun use of extra pictures - include only those figures essential to presenting results

### **Title Page:**

Choose a revealing title. It should be short. It should not have non-standard acronyms or abbreviations. It should not exceed two printed lines. It should include the name(s) and address (es) of all authors.



## Abstract:

The summary should be two hundred words or less. It should briefly and clearly explain the key findings reported in the manuscript-- must have precise statistics. It should not have abnormal acronyms or abbreviations. It should be logical in itself. Shun citing references at this point.

An abstract is a brief distinct paragraph summary of finished work or work in development. In a minute or less a reviewer can be taught the foundation behind the study, common approach to the problem, relevant results, and significant conclusions or new questions.

Write your summary when your paper is completed because how can you write the summary of anything which is not yet written? Wealth of terminology is very essential in abstract. Yet, use comprehensive sentences and do not let go readability for briefness. You can maintain it succinct by phrasing sentences so that they provide more than lone rationale. The author can at this moment go straight to shortening the outcome. Sum up the study, with the subsequent elements in any summary. Try to maintain the initial two items to no more than one ruling each.

- Reason of the study - theory, overall issue, purpose
- Fundamental goal
- To the point depiction of the research
- Consequences, including definite statistics - if the consequences are quantitative in nature, account quantitative data; results of any numerical analysis should be reported
- Significant conclusions or questions that track from the research(es)

## Approach:

- Single section, and succinct
- As an outline of job done, it is always written in past tense
- A conceptual should situate on its own, and not submit to any other part of the paper such as a form or table
- Center on shortening results - bound background information to a verdict or two, if completely necessary
- What you account in an conceptual must be regular with what you reported in the manuscript
- Exact spelling, clearness of sentences and phrases, and appropriate reporting of quantities (proper units, important statistics) are just as significant in an abstract as they are anywhere else

## Introduction:

The **Introduction** should "introduce" the manuscript. The reviewer should be presented with sufficient background information to be capable to comprehend and calculate the purpose of your study without having to submit to other works. The basis for the study should be offered. Give most important references but shun difficult to make a comprehensive appraisal of the topic. In the introduction, describe the problem visibly. If the problem is not acknowledged in a logical, reasonable way, the reviewer will have no attention in your result. Speak in common terms about techniques used to explain the problem, if needed, but do not present any particulars about the protocols here. Following approach can create a valuable beginning:

- Explain the value (significance) of the study
- Shield the model - why did you employ this particular system or method? What is its compensation? You strength remark on its appropriateness from a abstract point of vision as well as point out sensible reasons for using it.
- Present a justification. Status your particular theory (es) or aim(s), and describe the logic that led you to choose them.
- Very for a short time explain the tentative propose and how it skilled the declared objectives.

## Approach:

- Use past tense except for when referring to recognized facts. After all, the manuscript will be submitted after the entire job is done.
- Sort out your thoughts; manufacture one key point with every section. If you make the four points listed above, you will need a least of four paragraphs.



- Present surroundings information only as desirable in order hold up a situation. The reviewer does not desire to read the whole thing you know about a topic.
- Shape the theory/purpose specifically - do not take a broad view.
- As always, give awareness to spelling, simplicity and correctness of sentences and phrases.

#### **Procedures (Methods and Materials):**

This part is supposed to be the easiest to carve if you have good skills. A sound written Procedures segment allows a capable scientist to replacement your results. Present precise information about your supplies. The suppliers and clarity of reagents can be helpful bits of information. Present methods in sequential order but linked methodologies can be grouped as a segment. Be concise when relating the protocols. Attempt for the least amount of information that would permit another capable scientist to spare your outcome but be cautious that vital information is integrated. The use of subheadings is suggested and ought to be synchronized with the results section. When a technique is used that has been well described in another object, mention the specific item describing a way but draw the basic principle while stating the situation. The purpose is to text all particular resources and broad procedures, so that another person may use some or all of the methods in one more study or referee the scientific value of your work. It is not to be a step by step report of the whole thing you did, nor is a methods section a set of orders.

#### **Materials:**

- Explain materials individually only if the study is so complex that it saves liberty this way.
- Embrace particular materials, and any tools or provisions that are not frequently found in laboratories.
- Do not take in frequently found.
- If use of a definite type of tools.
- Materials may be reported in a part section or else they may be recognized along with your measures.

#### **Methods:**

- Report the method (not particulars of each process that engaged the same methodology)
- Describe the method entirely
- To be succinct, present methods under headings dedicated to specific dealings or groups of measures
- Simplify - details how procedures were completed not how they were exclusively performed on a particular day.
- If well known procedures were used, account the procedure by name, possibly with reference, and that's all.

#### **Approach:**

- It is embarrassed or not possible to use vigorous voice when documenting methods with no using first person, which would focus the reviewer's interest on the researcher rather than the job. As a result when script up the methods most authors use third person passive voice.
- Use standard style in this and in every other part of the paper - avoid familiar lists, and use full sentences.

#### **What to keep away from**

- Resources and methods are not a set of information.
- Skip all descriptive information and surroundings - save it for the argument.
- Leave out information that is immaterial to a third party.

#### **Results:**

The principle of a results segment is to present and demonstrate your conclusion. Create this part a entirely objective details of the outcome, and save all understanding for the discussion.

The page length of this segment is set by the sum and types of data to be reported. Carry on to be to the point, by means of statistics and tables, if suitable, to present consequences most efficiently. You must obviously differentiate material that would usually be incorporated in a study editorial from any unprocessed data or additional appendix matter that would not be available. In fact, such matter should not be submitted at all except requested by the instructor.



## Content

- Sum up your conclusion in text and demonstrate them, if suitable, with figures and tables.
- In manuscript, explain each of your consequences, point the reader to remarks that are most appropriate.
- Present a background, such as by describing the question that was addressed by creation an exacting study.
- Explain results of control experiments and comprise remarks that are not accessible in a prescribed figure or table, if appropriate.
- Examine your data, then prepare the analyzed (transformed) data in the form of a figure (graph), table, or in manuscript form.

### What to stay away from

- Do not discuss or infer your outcome, report surroundings information, or try to explain anything.
- Not at all, take in raw data or intermediate calculations in a research manuscript.
- Do not present the similar data more than once.
- Manuscript should complement any figures or tables, not duplicate the identical information.
- Never confuse figures with tables - there is a difference.

### Approach

- As forever, use past tense when you submit to your results, and put the whole thing in a reasonable order.
- Put figures and tables, appropriately numbered, in order at the end of the report
- If you desire, you may place your figures and tables properly within the text of your results part.

### Figures and tables

- If you put figures and tables at the end of the details, make certain that they are visibly distinguished from any attach appendix materials, such as raw facts
- Despite of position, each figure must be numbered one after the other and complete with subtitle
- In spite of position, each table must be titled, numbered one after the other and complete with heading
- All figure and table must be adequately complete that it could situate on its own, divide from text

### Discussion:

The Discussion is expected the trickiest segment to write and describe. A lot of papers submitted for journal are discarded based on problems with the Discussion. There is no head of state for how long a argument should be. Position your understanding of the outcome visibly to lead the reviewer through your conclusions, and then finish the paper with a summing up of the implication of the study. The purpose here is to offer an understanding of your results and hold up for all of your conclusions, using facts from your research and generally accepted information, if suitable. The implication of result should be visibly described. Infer your data in the conversation in suitable depth. This means that when you clarify an observable fact you must explain mechanisms that may account for the observation. If your results vary from your prospect, make clear why that may have happened. If your results agree, then explain the theory that the proof supported. It is never suitable to just state that the data approved with prospect, and let it drop at that.

- Make a decision if each premise is supported, discarded, or if you cannot make a conclusion with assurance. Do not just dismiss a study or part of a study as "uncertain."
- Research papers are not acknowledged if the work is imperfect. Draw what conclusions you can based upon the results that you have, and take care of the study as a finished work
- You may propose future guidelines, such as how the experiment might be personalized to accomplish a new idea.
- Give details all of your remarks as much as possible, focus on mechanisms.
- Make a decision if the tentative design sufficiently addressed the theory, and whether or not it was correctly restricted.
- Try to present substitute explanations if sensible alternatives be present.
- One research will not counter an overall question, so maintain the large picture in mind, where do you go next? The best studies unlock new avenues of study. What questions remain?
- Recommendations for detailed papers will offer supplementary suggestions.

### Approach:

- When you refer to information, differentiate data generated by your own studies from available information
- Submit to work done by specific persons (including you) in past tense.
- Submit to generally acknowledged facts and main beliefs in present tense.





ADMINISTRATION RULES LISTED BEFORE  
SUBMITTING YOUR RESEARCH PAPER TO GLOBAL JOURNALS INC. (US)

Please carefully note down following rules and regulation before submitting your Research Paper to Global Journals Inc. (US):

**Segment Draft and Final Research Paper:** You have to strictly follow the template of research paper. If it is not done your paper may get rejected.

- The **major constraint** is that you must independently make all content, tables, graphs, and facts that are offered in the paper. You must write each part of the paper wholly on your own. The Peer-reviewers need to identify your own perceptives of the concepts in your own terms. NEVER extract straight from any foundation, and never rephrase someone else's analysis.
- Do not give permission to anyone else to "PROOFREAD" your manuscript.
- **Methods to avoid Plagiarism is applied by us on every paper, if found guilty, you will be blacklisted by all of our collaborated research groups, your institution will be informed for this and strict legal actions will be taken immediately.)**
- To guard yourself and others from possible illegal use please do not permit anyone right to use to your paper and files.



CRITERION FOR GRADING A RESEARCH PAPER (COMPILATION)  
BY GLOBAL JOURNALS INC. (US)

Please note that following table is only a Grading of "Paper Compilation" and not on "Performed/Stated Research" whose grading solely depends on Individual Assigned Peer Reviewer and Editorial Board Member. These can be available only on request and after decision of Paper. This report will be the property of Global Journals Inc. (US).

Topics	Grades		
	A-B	C-D	E-F
<i>Abstract</i>	Clear and concise with appropriate content, Correct format. 200 words or below	Unclear summary and no specific data, Incorrect form  Above 200 words	No specific data with ambiguous information  Above 250 words
<i>Introduction</i>	Containing all background details with clear goal and appropriate details, flow specification, no grammar and spelling mistake, well organized sentence and paragraph, reference cited	Unclear and confusing data, appropriate format, grammar and spelling errors with unorganized matter	Out of place depth and content, hazy format
<i>Methods and Procedures</i>	Clear and to the point with well arranged paragraph, precision and accuracy of facts and figures, well organized subheads	Difficult to comprehend with embarrassed text, too much explanation but completed	Incorrect and unorganized structure with hazy meaning
<i>Result</i>	Well organized, Clear and specific, Correct units with precision, correct data, well structuring of paragraph, no grammar and spelling mistake	Complete and embarrassed text, difficult to comprehend	Irregular format with wrong facts and figures
<i>Discussion</i>	Well organized, meaningful specification, sound conclusion, logical and concise explanation, highly structured paragraph reference cited	Wordy, unclear conclusion, spurious	Conclusion is not cited, unorganized, difficult to comprehend
<i>References</i>	Complete and correct format, well organized	Beside the point, Incomplete	Wrong format and structuring



# INDEX

---

---

## **A**

Aforementioned · 19  
Agora · 15, 16, 19, 44  
Artefacts · 10

---

## **B**

Bayesian · 45, 47, 53, 74, 78, Lxxxi

---

## **C**

Chemises · 13  
Cloudmade · 60, 62  
Colossal · 58  
Conjectured · 36  
Conjectures · 23  
Constructivism · 5, 6, 11, 39

---

## **D**

Demarcates · 35  
Didactic · 2, 23, 35, 41  
Dilate · 30

---

## **E**

Elicited · 32  
Emanating · 2  
Epigrammatically · 4  
Euclidean · 6, 8, 24, 45, 55, 66, 67, 69, 78  
Excavation · 15

---

## **F**

Fractal · 2, 5, 12, 23, 25, 26, 30, 31, 65

---

## **H**

Henceforth · 12  
Heuristics · 5  
Hollebrands · 9, 10, 40

---

## **I**

Infinitesimals · 26

---

## **M**

Mandelbrot · 25, 26  
Monocotyledoneae · 51  
Mosaics · 16

---

## **P**

Perceptron · 45

---

## **R**

Rhomboid · 12

---

## **S**

Sierpinski · 26, 28

---

## **T**

Trilateration · 58

---

## **V**

Verbalizing · 11  
Visuospatial · 17  
Vygotsky · 6, 8, 44

---



save our planet



# Global Journal of Computer Science and Technology

Visit us on the Web at [www.GlobalJournals.org](http://www.GlobalJournals.org) | [www.ComputerResearch.org](http://www.ComputerResearch.org)  
or email us at [helpdesk@globaljournals.org](mailto:helpdesk@globaljournals.org)



ISSN 9754350