Data-Driven Knowledge Agriculture: A Paradigm Shift for Enhancing Farm Productivity & Global Food Security

By Aprajita Srivastava & Dr. H. O. Srivastava

World Development Foundation

Abstract- Data-driven Knowledge agriculture using mechanized intelligent computer-based monitoring and control systems and complex Software for machine learning and visualization for predicting a variety of parameters such as future food requirements, resource planning for higher yield, and supply chain is the future of farming. This needs to be urgently adopted by the world farming community to provide food to the growing world population, remove hunger, and at the same time sustain planet resources by judicious uses of input such as water, fertilizer, pesticide etc., as envisioned by Sustainable Development Goals 2030. This paper discusses data-driven technology for identifying trends and other insights for making informed decisions for enhanced productivity and profitability, through market research and evaluating customer needs and sentiments.

Keywords: precision agriculture, knowledge agriculture, data-driven agriculture, SDG2030, farm data visualization, agriculture analytics, agriculture business intelligence.

GJSFR-A Classification: FOR Code: 070199

DATADRIVENKNOWLEDGEAGRICULTUREAPARADIGMSHIFTFORENHANCINGFARMPRODUCTIVITYGLOBALFOODSECURITY

Strictly as per the compliance and regulations of:

Data-Driven Knowledge Agriculture: A Paradigm Shift for Enhancing Farm Productivity & Global Food Security

Aprajita Srivastava [α] & Dr. H. O. Srivastava [σ]

Abstract- Data-driven Knowledge agriculture using mechanized intelligent computer-based monitoring and control systems and complex Software for machine learning and visualization for predicting a variety of parameters such as future food requirements, resource planning for higher yield, and supply chain is the future of farming. This needs to be urgently adopted by the world farming community to provide food to the growing world population, remove hunger, and at the same time sustain planet resources by judicious uses of input such as water, fertilizer, pesticide etc., as envisioned by Sustainable Development Goals 2030. This paper discusses data-driven technology for identifying trends and other insights for making informed decisions for enhanced productivity and profitability, through market research and evaluating customer needs and sentiments.

Keywords: precision agriculture, knowledge agriculture, data-driven agriculture, SDG2030, farm data visualization, agriculture analytics, agriculture business intelligence.

I. Introduction

Knowledge Agriculture™ is a new way of farming that uses technology and tools such as IOT (Internet of Things), Robotics, AI (Artificial Intelligence), UAV (Unmanned Aerial Vehicles), Cloud, Greenhouse, BA (Business Analytics)and specialized software for weather modeling, smart zone seeding, fertilizer modeling, *(Srivastava, 2018; Definitions,2022),* to address issues relating to food requirements of current 8.0 billion world population, estimated to grow to 9.8 billion in 2050 *(PRB, 2016)*. The current practice of farming in most of the countries is based on traditional methods of sowing, harvesting, storage and marketing which results in poor productivity and profit. We define agricultural productivity as the ratio of outputs to inputs, expressed either in volumes or in physical quantities (kg, tons, etc.) *(FAO, 2018)*. TFP (Total Factor Productivity) is a measure of the efficiency of the contribution of all the significant inputs into production. TFP provides a complete picture of productivity and is more closely connected to unit production costs and market prices than partial productivity indicators *(Fuglie, 2015)*. There is a close relationship between agricultural productivity and farm incomes. Considering this, increasing farm income is at the center of measures related to food security, rural livelihoods, ending hunger in Africa and other parts of the Globe by 2030 and fulfilling SDG2030 targets on food and nutrition security. Despite the importance of agricultural productivity, data on various parameters required to analyze and deduce the current productivity and forecast the model for improved productivity is scarce and of poor quality. There is a need for new and enhanced data collection frameworks to better measure agricultural production and the amounts of inputs used in the production processes.

Data is becoming an ever-important factor in the world, and businesses are developing the skill to use data analysis as a tool for enhanced profit. Agriculture is not immune to this, and data-driven decisions are essential for improved agricultural productivity. Data analysis focuses on cleaning, modeling and visualizing data to provide descriptive, diagnostic, predictive, prescriptive and cognitive analytics. Descriptive analytics based on historical data may be used to generate reports to give a view of a farmer's production, sales and financial data. Diagnostic analytics helps in answering questions about why certain events happened for example anomalies that might be due to unexpected changes in a metric or a particular market. Predictive analytics techniques use historical data to identify trends, for instance changes in demand or consumption and determine if they're likely to recur. Prescriptive analytics is an application of(ML) (Machine Learning)that prescribes optimal actions to achieve a goal or target, for example use of lesser pesticides due to consumer's behavior of changing to organic food. It may also help farmers provide historical crop yield record with a forecast reducing risk management. Machine learning uses statistical algorithms that involve statistical and functional analysis of existing data to learn, a process called training. Patterns and relationships in the data identified during training are used to build a model. The model makes intelligent decisions about data it hasn't encountered before. A process called inference is used to make decisions

Corresponding Author α: Cloud support engineer in analytics, and Hon. Director of World Development Foundation undertook R&D projects for use of ICT for agriculture in India and Ethiopia for increased productivity and social projects for citizens to express themselves socially, culturally, politically and spiritually, Krishi Apts. Vikaspuri-D, New Delhi-110018, India. e-mail: aprajitaedu1311@gmail.com

Author σ: Ph. D., FIETE, Former Addl. Director General & Head of AIR Resources of All India Radio and Doordarshan (Level of Addl. Secretary to the Govt. of India), is the President and CEO of World Development Foundation.

about new data. Training ML models with larger volumes, more relevant, and accurate data increases the forecasting accuracy. Cognitive analytics helps analyse the various changes in a parameter, its effect, and its solution *(Asakiewicz, 2016)*. Extensive studies on the predictive ability of ML techniques, such as multiple linear regression, regression trees, artificial neural networks, support vector regression, and k-nearest neighbor for crop yield production, have been undertaken *(Gonzalez-Sanchez, 2014)*. ML and advanced sensing techniques have been used to analyze online multi-layer soil data, and satellite imagery of crop growth characteristics to predict wheat yield *(Pantazi et al, 2016)*. 'Crop Advisor' is a software tool for predicting the influence of climatic parameters on crop yields *(Veenadhari et al, 2014; Majumdar et al,2017)* SMAG, a French company has developed software using 30 years of weather data history, satellite and drone images, and soil types to enable users to track the progress during the life cycle of a plant and predict yields. 80% of French agricultural land under wheat cultivation is managed through this algorithm. InVivo, France's leading agricultural cooperative group with 220 members using the technology, earns €6.4 billion in sales *(Talend, n.d.)*.

II. Knowledge Agriculture™

Knowledge Agriculture aims to achieve the objectives enshrined in SDG2030 to increase the agriculture yield without harming the environment, sustaining planet resources, and simultaneously countering challenges such as depletion of water resources and increased erosion and loss of productivity due to the occurrence of extreme weather events. The recent global environmental changes are apparent caution to mankind to immediately change the way it produces food and makes it available to the present and future generations. This can be achieved by optimal use of input resources such as seed, water, fertilizer and chemicals by preventing loss in storage, transportation and supply. In knowledge agriculture, data is captured from several sources. Sensors and IOT (Internet of Things) are used to collect farm data such as soil moisture, humidity, temperature, heat, light, etc. A second database stores the survey report that includes, productivity of soils, suitability of soils for raising specific crops, etc. Another database contains details of time for planting, produce to be planted, row spacing, desired yield, waste recycling, water supply, etc. There are several other databases which store historical rates, market details, supply chains, etc. These structured, semi-structured, or unstructured data are ingested into local or cloud-based storage, computing, analyzing, and creating models for ML-based predictions and taking corrective actions. The data allows the farmers to use a DSS (Decision Support System), to make optimized decisions such as tending each plant with water and fertilizers, spraying pesticides, and eliminating the weeds by selective burning or tilling as per its need during the life cycle without damaging the desired plant. Further, supply chain network design, product design and development, demand planning, procurement management, customized production, inventory management, logistics, and agile supply chain are essential components in farm management. A sustainable global supply chain is becoming an ever-increasingly complex system, especially with the need to deal with international partners*(McCue, 2020).*Modern technologies, such as CEP(Complex Event Processing), RFID, block chain, IoT, and WSN(Wireless Sensor Networks), boost supply chain performance. Supply chain analytics allows quick adjustments and more effective tactical decisions.

III. Storytelling from Farm Data

a) *System Architecture for the use of machine learning models for Knowledge Agriculture*

A system architecture for predictive analysis of historical and live data and taking corrective actions is shown in Figure 1.
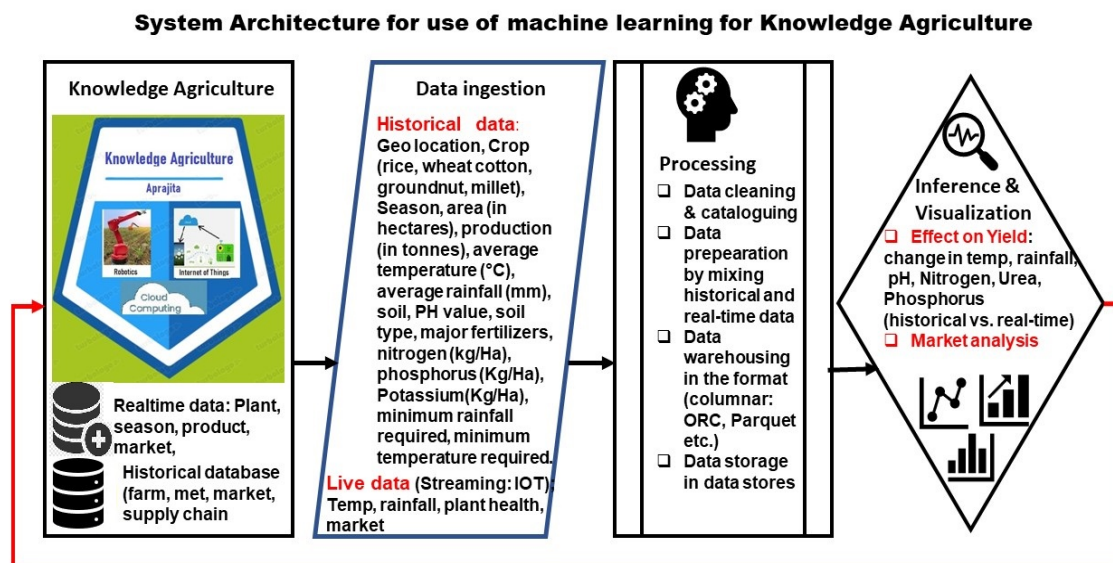
System Architecture for use of machine learning for Knowledge Agriculture

Figure 1: System architecture for predictive analysis of historical and live data and taking corrective actions

From the Figure, it will be seen that Knowledge Agriculture involves the use of IOT (Internet of Things) tools such as infra-red cameras, sensors (barometers, hygrometers, light meters, and ambient thermometers), electrodes, GPS (Global Positioning System) guided farm machines (tractor, drones) to detect physical parameters namely moisture, heat, light, pressure, nutrients and pests, etc. These data are streamed directly to the cloud, the technology that relieves the users from maintaining an elaborate computer system with expert staff including programmers and system administrators, for processing, monitoring, and analysis. The historical database is also held on Geo location of the farm, type of the crop (rice, wheat cotton, groundnut, millet), seasons when they are sowed/harvested, area (in hectares), production (in tonnes), average temperature (℃), average rainfall (mm), soil, pH (Hydrogen potentials) value, soil type, major fertilizers, nitrogen (kg/Ha), phosphorus (Kg/Ha), Potassium (Kg/Ha), minimum rainfall required, the minimum temperature required, etc. for optimum yield. Data on logistics, marketing chain, price (historical and current), consumers and input suppliers are also maintained on the cloud. Meteorological data is used for analyzing the weather variation. Bringing all the data to the cloud is part of data ingestion.

The data coming from different agencies may be in structured, unstructured, or in semi-structured formats. These need to be cleaned, normalized, categorized, and cataloged. Since the data is enormous and petabytes of big data need to be analyzed, it needs to be put in a way that analysis becomes possible, a few in real time for taking immediate corrective action in the farm such as tending each plant with water, fertilizers as per its need during the life cycle; eliminating the weeds

by the spraying of pesticides, tilling or selective burning without damaging the desired plant. It may need certain data available in CSV (Comma-Separated Values) or JSON (JavaScript Object Notation) formats to be converted in a columnar form known as ORC or Parquet form. These activities come under processing.

The final analysis is done by building visualization for real-time monitoring of the crop for dry fields or identifying plants needing attention for the use of fertilizer, pesticide or heating, etc., and thus allocating resources dynamically and efficiently. Careful and precise application of all inputs for agricultural produce can maximize crop yields and thereby reducing cost, and at the same time minimize waste of pesticide and water usage. Dynamic evaluation of moisture content can stop the excessive use of water for irrigation. This is especially of great value for 18 countries, with almost four billion people who are using their water reserves to the point of exhaustion to produce grains*(Earth Policy Institute, 2013).*The consolidation of historical weather data with meteorological prediction can be used to build complex models to predict the time of sowing/harvesting the crop. Another real-time model for supply chain can track the movement of food to identify bottlenecks and optimize the movement of product by improving logistics. To summarize, the analysis provides the potential to foresee demand, assess user sentiment, promote coordination and collaboration throughout the supply chain through fleet management tracking solutions, and use GPS-oriented analytics to optimize routing, cut transportation costs and provide an advanced mapping of vehicle locations.

Big data analytics is a complex process of examining the massive amount of data to uncover information such as hidden patterns, correlations, and

other insights like market trends and customer sentiments that enables organizations to make informed business decisions. We divide the entire analysis in three steps, namely Data management, Analytics & visualization. This is shown in Figure 2. A few tools used in different stages are mentioned in the Figure.
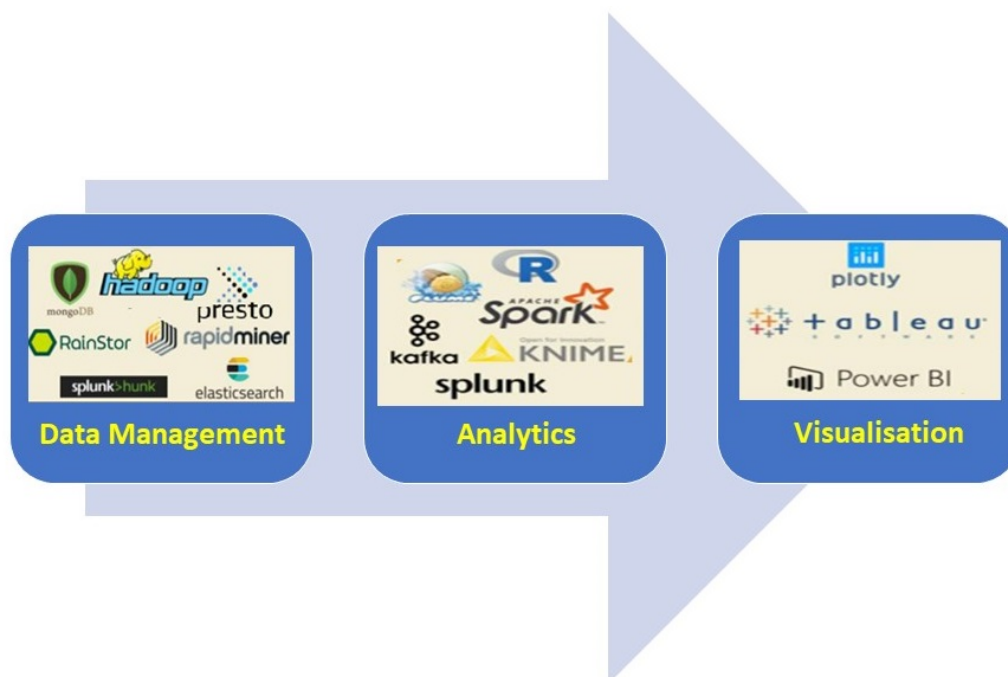

Figure 2: Stages of Data analytics and visualization

Hadoop *(Apache Hadoop Documentation, 2014)*is an open-source framework written in Java that provides many analytical tools to store and process large datasets ranging in size from gigabytes to petabytes to generate new insight, which includes Machine Learning and data mining. The Apache Hadoop kernel has a storage part, known as Hadoop Distributed File System (HDFS), and a processing part known as the Map Reduce algorithm for parallel processing. Few of the tools are Apache spark *(Spark Core Programming, n.a.; Kannan, 2015)*, a cluster computing platform to process batch applications, Machine Learning, streaming data processing, and interactive queries; Map Reduce *(Seema & Jha, 2015)*, an Algorithm based on the YARN framework to perform the distributed processing in parallel in a Hadoop cluster; Apache Hive *(Hiba et al. 2019)*, a Data warehousing tool that uses query language known as HQL or HIVEQL; Apache Impala, an open-source SQL engine; Apache Mahout*(Anil et al. 2020)*, used for implementing various Machine Learning algorithms offering implementations of classification, clustering, dimensionality reduction and linear algebraic computations; Apache pig, an open-source Apache library that runs on top of Hadoop, providing a scripting language used for analysing massive datasets by representing them as dataflow *(Swarna & Zahid 2017)*; HBase, a non-relational, NoSQL distributed, and column-oriented database that allows for data to be analysed in real-time, as it is entered (History of Apache HBase, n. a.); Tableau, a software to generate helpful visualizing charts on interactive dashboards and worksheets and many more used in the Business Intelligence Industry *(Tableau, n. a.)*. These tools, with their short descriptions, have been shown in Figure 3.
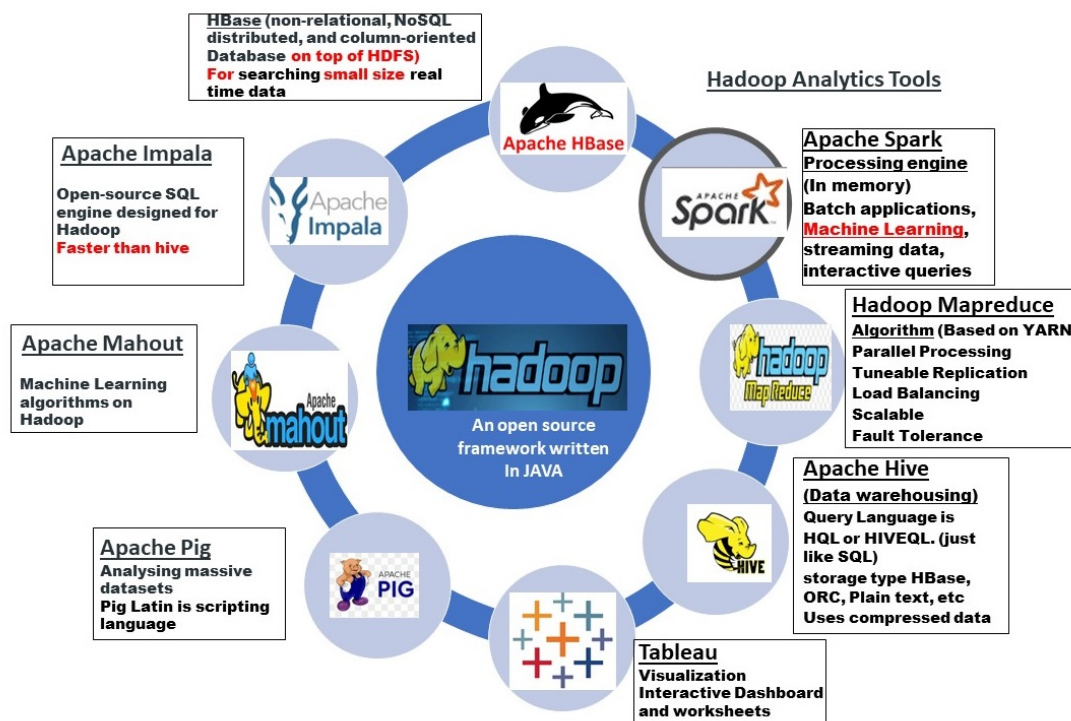
Figure 3: Hadoop Analytics tools for Big Data analysis

IV. Cloud-Based Architecture

Open-source tools can be used for Knowledge Agriculture data analytics, visualization, and undertaking corrective action using BI (Business Intelligence) tools like a tableau. However, this needs hardware and software infrastructure with competent programmers. The alternate solution is to use Cloud-Based platform which offers "SaaS (Software as a Service), PaaS (Platform as a Service), and IaaS (Infrastructure as a Service)"*(IBM Cloud Education, n. a.)*. SaaS utilizes the internet to deliver applications that run directly through web browsers, thus sparing the employees and companies from the efforts of installing, managing, and upgrading software. PaaS provides a framework for developers to build upon and create customized applications. IaaS is a form of virtual cloud computing technology that provides highly scalable infrastructure, including servers, networks, operating systems, and storage. Using IaaS, businesses hire resources on-demand and scale it as per their requirements instead of having to buy the hardware outright. Google Cloud Platform Services (GCP), Amazon Web Services (AWS), Microsoft Azure, Digital Ocean, Linode, Rackspace, and Cisco Metapod are a few companies providing IaaS. In 2020, the cloud-based segment had the largest share of the agriculture analytics market. The large share of cloud-based market is primarily due to easy accessibility, lower maintenance, affordable pricing, and reliable security. Furthermore, benefits provided by cloud-based solutions in terms of SaaS, PaaS, and IaaS for agriculture analytics are helping the segment grow at the highest speed. A few basic architectures based on a few major cloud service providers for big farm data analytics are discussed below.

a) Google Cloud Platform (GCP)-based architecture

A typical system architecture based on the Google cloud platform is shown in Figure 4. Here, data from IOT devices are captured using MQTT protocol through IOT core (fully managed Service for securely connecting and managing millions of IoT devices), and data is prepared through DataPrep (explores, cleans, and prepares structured and unstructured data for analysis). Thereafter this data is analyzed along with the other historical database stored in Firestore (a NoSQL document database), using Cloud Functions (a serverless execution environment for building and connecting cloud services) and BigQuery/ML that creates and executes machine learning models. The information after analysis is made available to Data analysts/farmers through Pub/Sub, which provides messaging between applications.
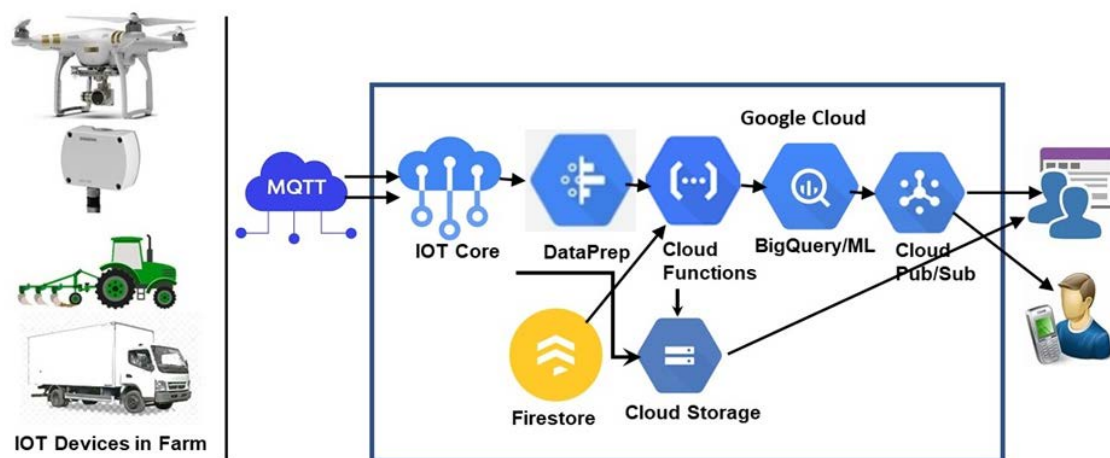
Figure 4: System architecture for farm data analysis and prediction based on Google cloud (Note: The GCP icons are available in the open domain)

i. *Use case-Taranis (Google cloud along with Tensor Flow)*

As much as forty percent of crops in agriculture are regularly lost due to pests, crop illnesses, weeds, and lack of nutrients. Using drone technology and artificial intelligence, Taranis *(Taranis, n.d.)*, working in several countries worldwide, with more than 20 million acres of land in Russia, Eastern Europe, and South America, is providing farmers with modern tools on the cloud along with Tensor Flow to reduce crop loss, increase yields, and lower costs *(Google Cloud, n.d.)*.To develop BI (Business Intelligence) models, drones were deployed to capture tens of millions of photographs over the past year and a half, which have been analyzed and tagged. Each photo has up to a thousand items of interest, such as insect damage or leaf discoloration. The company claims to have processed around 100 million distinct features in approximately 700,000 images. The insights provided through visualization provide farmers with the information that enables them to intervene early and prevent crop loss. Farmers can target problems with concrete solutions, like adding fertilizer in a specific area with low in nutrients. The scalable Compute Engine dedicates information from the images when new images arrive. Data is uploaded to a Cloud SQL database for further analysis.

Tensor Flow *(TensorFlow, n. a.)*is an *open-sourced* machine learning platform that delivers a complete functional solution with a focus on deep neural networks to analyze massive amounts of unstructured data. It contains a diverse and complete set of libraries, tools, and community resources. Tensor Flow is used for building the machine learning models by using tens of millions of photographs of agriculture fields for training. Feeding the dataset for the new photographs and comparing them with models is used to discover items of interest, such as insect damage or leaf discoloration.

b) *Amazon Web Services (AWS) based architecture*

A typical broad system architecture using AWS is shown in Figure 5. It would be observed from the Figure that broad modalities are similar. The IOT data is streamed using a tool known as Kinesis which allows real-time streaming of IOT data from the devices deployed on the farm. Another tool, IoT Core connects billions of IoT devices such as sensors, actuators, embedded devices, wireless devices, and smart appliances over MQTT, HTTPS, and LoRa WAN protocols and routes trillions of messages to users without managing any infrastructure.
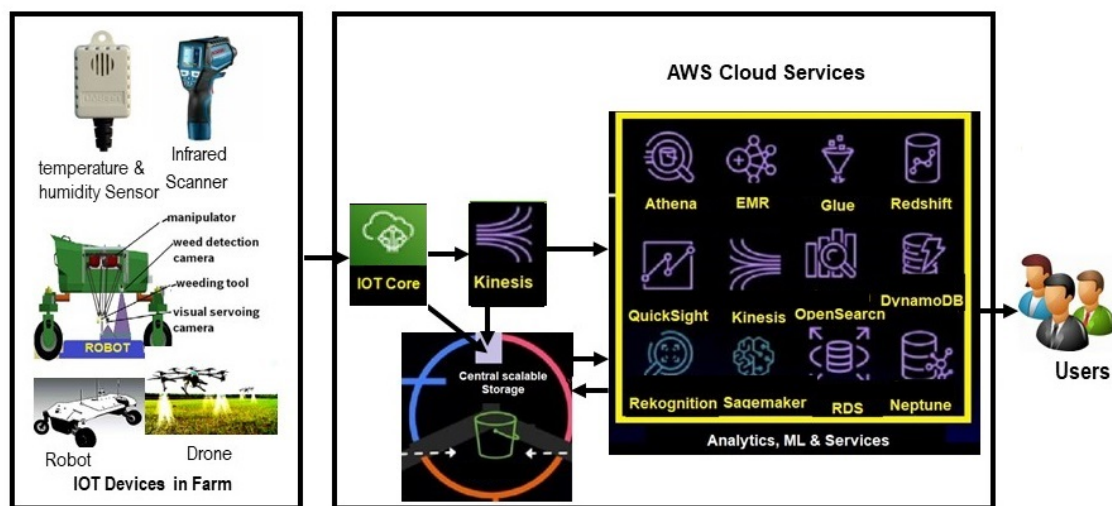
Figure 5: System architecture for farm data analysis and prediction based on Amazon Web Services (AWS) (Note: The AWS icons are available in the open domain)

Kinesis Data Analytics (KDA) analyzes streaming data to gain actionable insights and responds to farmers' needs in real-time. It also generates real-time alerts and dashboards. KDA supports standard ANSI SQL and provides an easy-to-use schema editor. It includes open-source libraries based on Apache Flink. IoT Analytics datasets can be taken to the Quick Sight console (Business intelligence tool) for creating dashboards and visualizations. Red shift is a high-performance, petabyte-scale data warehouse service for analytical processing (OLAP) workloads, while RDS Postgre SQL or Aurora typically stores terabytes of relational data for on-line transaction processing (OLTP) workloads. Dynamo Db stores non-relational data, and Neptune is a fully managed graph database. Glue is a server less data integration and ETL (Extract, Transform and Load) tool used for discovering, preparing, and combining data for data analysis and Machine Learning. Open Search provides for free-text search. Athena is a serverless interactive query service, an upgraded SQL query engine from the Trino *open-source* project, that makes it easy to analyze data in storage using standard SQL. Amazon has developed their own EMR, which is taking the place of Apache Hadoop and Apache Spark. This makes it simpler to perform tasks that involve big data analysis such as data transformation, graph analytics, and machine learning on a managed cluster platform. Rekognition is part of the artificial intelligence services provided for image analysis.

i. *Use cases*

Mantle Labs*(Mantle Labs, n. a.)*uses the AWS cloud platform to automate the processing of satellite data from 500 million hectares of agricultural land spanning six continents for tracking vegetation health, soil moisture, and climate impact daily. This service provides an accurate, near real-time risk assessment to the world's 600m farmers, instantly. The company states

that it can accumulate and analyze billions of pixels of imagery from several satellites on a daily basis to create a data cube of the entire world. It uses Elastic Container Registry (ECR) for storing container images and Fargate as a serverless compute engine. ECR allows sharing and downloading of images securely over Hypertext Transfer Protocol Secure (HTTPS) with automatic encryption and access controls. Fargate is a serverless compute engine that enables building applications without managing servers. Mantle Labs also leverages Simple Queue Service (SQS) for event-based batch processing of satellite images. The data analysis forecasts crop health, forest changes, climate impact, crop observation, drought impact analysis, and field-level profitability analysis. The deep learning engine has the capability to examine petabyte-scale data to provide high-resolution analysis for any area on the planet. The farmers can use the platform for crop monitoring and near real-time production estimation, including commodity price forecasts, production, and supply chain analysis. To bank and insurance companies, it offers credit risk assessment, underwriting and pricing, portfolio creation and risk diversification, claims fraud management, and processing *(AWS, n. a.).* Mantle Labs, with the Agri-EPI Centre, offered its Geobotanics platform, which runs crop monitoring solutions, for UK (United Kingdom) food security during the coronavirus pandemic ("*Crops and Soil Newson Agri-tech partners support COVID-19 food security*", 2020).

GeoPard, Köln, Germany *(GeoPard, n. a.)* uses AWS for spatial data analytics and artificial intelligence algorithms for crop farming that automates workflows to create VRA maps for seeding, fertilizing, and crop protection based on various data sources and agronomic logic. Variable-rate technology (VRT) allows determination of the exact number of passes over an area to deliver the right amount of farm inputs such as

24

fertilizer, chemicals, pesticide, lime, gypsum, and irrigation water. By creating a Zones Map from specified data layers such as Topography, Satellite Images, Soil Sampling, High-Density Sensors and Cameras, Applied Datasets of Yield, VRA maps can be employed for any agricultural field. Data visualization leads to new ideas, and analytics helps to solve problems by creating solutions and writing prescriptions agronomists.

CropX captures raw data from global soil sensors and integrates it with topography/soil maps, weather forecasts data, satellite imagery, hydraulic models, crop models, user inputs, data from the controllers, and agricultural equipment such as sprayers, tractors & harvesters, and then ingests it to a centralized platform running on Elastic Compute Cloud (Amazon EC2) instances. It then analyses and saves satellite imagery, agronomical data, and insights in the cloud storage. It uses SNS to transfer messages between services. The Service provides crop-specific and growth-stage-specific insights and recommendations, such as advising farmers when, where, and how much to irrigate, fertilize and spray *(CropX, n. a.)*.

c) Microsoft Azure

Farm Vibes. AI *(Project Farm Vibes, n.a.)* is Microsoft's *open-source* Multi-Modal Geospatial ML Models for Agriculture and Sustainability. There are three main components of FarmVibes.AI. The first one consists of data ingestion from sources such as satellite imagery (RGB, SAR, multispectral), drone imagery, weather data, custom sensor data (such as weather sensors), and more. The input data is pre-processed for building ML models with ease based on parameters that can be specified. The inbuilt notebooks are used to tune the models to achieve a level of accuracy for the specific parts of the world or seasons. The library includes data for detecting practices (for example, harvest date detection in the country for a particular crop), estimating climate impact (both seasonal carbon footprint and long-term sustainability), microclimate prediction, and crop identification. The artificial intelligence arm of Farm Vibes has three parts, namely, Async Fusion, which pairs sensors and satellite imagery to create nutrient heat and soil moisture maps; Space Eye, which removes clouds from satellite imagery; Deep MC, which uses sensor data and weather forecasts to predict the temperatures and wind speeds of a farm's microclimate; and a "what if" analysis tool that predicts effect on the soil's carbon sequestration abilities based on variances in farming methods.

i. Use Case-Bayer

Bayer's Climate Field View™ is a data integration tool that provides yield analysis, field region reports, field health imagery, and manual seed scripts. As per the company's report, the Software is being used on more than 180 million farming acres across more than

20 countries. Microsoft and Bayer have entered into a strategic partnership to build a new cloud-based set of digital tools and data science solutions to enable the rapid development of agriculture food tech applications by the combination of modern cloud technology and adding a layer of data *(Digital Farming, n. a.)*.

V. Challenges

Soil fertility has severely degraded in many parts of the world due to intensive agriculture, increasing use of fertilizers and pesticides, over-grazing, water pollution, deforestation, salinization, and accumulation of non-biodegradable waste. Climate change due to global warming further aggravates land degradation, soil erosion, and soil fertility. Natural disasters like flooding and landslides are being witnessed more frequently worldwide. Severe soil degradation can reduce crop yield by over 50%*(FAO, 2019)*. At the same time, Food and Agriculture Organization of the United Nations has estimated that the world population will exceed 9 billion by 2050, and the world will lose about 250 million crop-production acres due to urbanization and soil degradation from excessive tillage and other farming practices *(Wilde, 2021)*. Food loss and waste (FLW) in production, postharvest and consumption stages are 24, 24, and 35 percent, respectively, accounting for more than 80 per cent of food wastage in these stages, which is quite alarming *(Li et al, 2017)*.

The agriculture analytics market is expected to grow to $2.27 billion by 2027, at a CAGR of 17.5% from 2020 to 2027. The key factors driving the growth are rising pressure to meet global demand for food, especially in the current world environment of food shortage due to the COVID epidemic, climate changes, and war. This environment has led to an increasing need for improved farm productivity that is possible only by the use of Knowledge Agriculture deploying modern tools and technology of AI and predictive analytics to improve the yield and cut down the losses of 1.2 billion tonnes globally; 40% of all food is never eaten when both farming and post-farming are taken into account, as per WWF-UK *(WWF, 2021)*. Going with cloud-based technologies and applying big data in the agribusiness chain will reverse the trend of farmers selling their land for choosing employment in the service sector and make the agriculture business profitable.

The lack of technical know-how among farmers, the fragmented agriculture industry, and need for heavy capital investment for Knowledge Agriculture, lack of standardization for data management and data aggregation may hamper the growth of this market. Further, political compulsions, farmers', and trade unions' lack of trust in industrialists and solution providers are significant impediments. One of the most significant challenges is discovering stable business and operational frameworks for private-public

collaborations. The Indian Government's withdrawal of farm bills is a pointer *(Mukherji in Business Standard, Jan, 17, 2023)*. The explicit motive for data harvesting is the prediction of consumers' wants and needs. The farmers fear that the data collected by service providers can be manipulated by them for their own advantages. For example, the forecast of a higher need for fertilizer in a specific region of the world may lead to an artificial increase in prices in that region in collusion of service provider and fertilizer supplier. Beyond data access and infrastructure, digital democracy calls for a fundamental redistribution of decision-making power from a small number of corporate stakeholders to a broader group of farmers and citizens. GODAN (Global Open Data for Agriculture and Nutrition) of United Nations, currently with more than 374 partners from national governments, non-governmental organizations, international and private sector organizations, pleads for the active sharing of data on agriculture and nutrition among all the stakeholders *(GODAN, n. a.)*. The European Commission's launch of public consultations on the Data Act with the possibility of imposing binding rules to provide governments access to private-sector data has started heated debates. The Data Act is a proposal by the European Commission for a legislative provision that "aims to create a fair data economy by ensuring access to and use of data for legitimate purposes, including in business-to-business and business-to-government [B2G] situations" (European Commission, 2021).

VI. Way-Ahead

Linking data from farm operations on agricultural sites with information and knowledge collected from laboratories will be required to be effectively analyzed and used to optimize sector performance. The best possible solution for this is to make accessible global data from land globally and provide a platform for analysis and use for long-term sustainable development by improving economic opportunities for farmers. Upgrading the skills of all the stakeholders, resource persons in agriculture Universities, students and farmers, open access to research, meta-analysis, and available data publication is vital for food nutritional security. Providing more specialized Big Data analysis tools on mobile devices, and extending the information about the features, benefits, and limitations of each framework in various local languages will help the world population. Carnegie Mellon University is undertaking a research project for automatic speech recognition for 2000 languages from the present 200 *(Science Daily, 2023)*. WDF University for SDG2030 has developed skill development on-line courses, which provide choice for translation of courses in many Indian and foreign languages *(WDF University, n. a.)*. Hence, technological issues related to ingestion, analysis, processing, and application associated with Farm Data must be included in studies and constitute an essential research topic.

VII. Conclusion

There is an urgent need to create a sustainable food system to feed the growing population on the earth. The climate crisis grew increasingly severe and irreversible in the year 2022. The world witnessed record ice melt at the poles, catastrophic flooding in Pakistan and China, record-breaking heat waves in the US (United States) and Europe, and severe drought in Africa. The war between Russia and Ukraine has led to rising energy prices and climate-fuelled droughts, and has pushed a perilously stretched global food system to the brink. The situation has led to an increased interest in investigating how our food is grown and ensuring that the most efficient practices are used. Efficiency in growing more food needs urgent use of tools and technology for Knowledge Agriculture. Countries have to adapt it to escape a cycle of catastrophes and debt. The 2030 agenda for 17 Sustainable Development Goals (SDGs) aims to end poverty and malnutrition around the globe, preserve the climate and environment while sustaining the planet's resources. The agenda was adopted in 2015, and now after, seven years (mid-term of 2015-2030), "the world continues to lose ground in its efforts to end hunger, food insecurity, and malnutrition", according to the report of SOFI(State of Food Security and Nutrition in the World) released in 2022 *(Food and Agriculture Organization of United Nations, 2022)*. The information admits that "many of the SDG2 targets are growing wider each year". Eight hundred twenty-eight million people, around 10.5% of the world population, were affected by hunger globally in 2021, an increase of 46 million since the end of 2020. Nearly 3.1 billion people globally could not afford a healthy diet in 2020, an increase of 112 million from 2019 due to the pandemic and war. The government of India has been providing free food, to prevent death from hunger, to its 814 million population since March 2020 under the National Food Security Act, which has been extended up to Dec. 2023 *(The Economic Time, 2022)*. The total adds to 3.924 billion of 7.79 billion, about 50% world population for the year 2020, lacking proper food during pandemic. Serious efforts are required to avoid the bleak future by using modern tools and technology to farm sustainably and climate sustainably and offer farmers an adequate advice, training, and skill.

References Références Referencias

1. Agri-tech partners support COVID-19 food security. (2020, April 21).*Crops and Soil News.*https://agri-epicentre.com/news/agri-tech-partners-support-covid19-food-security.

2. Anil, Robin & Capan, (2020). Apache Mahout: Machine Learning on Distributed Dataflow Systems. Journal of Machine Learning Research. 21. 1-6.

3. Apache Hadoop Documentation. (2014). http://hadoop.apache.org/

4. Asakiewicz, Christopher, Cognitive Analytics for Making Better Evidence-Based Decisions (August 24, 2016). Available at SSRN: https://ssrn.com/abstract=2965767 or http://dx.doi.org/10.2139/ssrn.2965767

5. AWS. Mantle Labs Improves Global Food Supply Chain Financing on AWS, https://aws.amazon.com/solutions/case-studies/mantle-labs/

6. Brown, L.R. (2013, July 09). Peak Water: What Happens When the Wells Go Dry? *Earth Policy Institute*. http://www.earth-policy.org/plan_b_updates/2013/update115

7. CropX Agronomic Farm Management System, https://cropx.com

8. Definitions.net. STANDS4 LLC.(2022). Knowledge agriculture. https://www.definitions.net/definition/knowledge+agriculture

9. Digital Farming, Smart Solutions for a Sustainable Future, Bayer. https://www.bayer.com/en/agriculture/digital-farming.

10. European Commission. (03 June 2021 - 03 September 2021). Data Act & amended rules on the legal protection of databases, https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13045-Data-Act-including-the-review-of-the-Directive-96-9-EC-on-the-legal-protection-of-databases-/public-consultation_en

11. Food and Agriculture Organization of United Nations (FAO). (2018). Guidelines for the measurement of productivity and efficiency in agriculture. https://www.fao.org/3/ca6395en/ca6395en.pdf

12. Food and Agriculture Organization of United Nations (FAO). (2019). Global Symposium on Soil Erosion (GSER19), https://www.fao.org/about/meetings/soil-erosion-symposium/key-messages/en/

13. Food and Agriculture Organization of UN. (2022). The State of Food Security and Nutrition in the World 2022. https://www.fao.org/documents/card/en/c/cc0639en

14. Fuglie, K., (2015). Accounting for growth in global agriculture. Bio-based and Applied Economics. 4(3), 201–234

15. GeoPard Agriculture, https://geopard.tech

16. GODAN. Global Partnership for Sustainable Development. Global Open Data for Agriculture and Nutrition. https://www.data4sdgs.org/partner/global-open-data-agriculture-and-nutrition

17. Gonzalez-Sanchez, Alberto., Frausto-Solis, Juan., Ojeda-Bustamante, W., (2014) Predictive ability of machine learning methods for massive crop yield prediction. Span J Agric Res. 12(2), 313–28.

18. Google Cloud: Helping farmers to feed the planet with cutting-edge drone imaging and AI, https://cloud.google.com/customers/taranis

19. Hiba. A., Abu-Alsaad, et. al. (2019). Retailing Analysis Using Hadoop and Apache Hive. https://ijssst.info/Vol-20/No-1/paper8.pdf

20. IaaS vs. PaaS vs. SaaS, IBM Cloud Education, https://www.ibm.com/in-en/cloud/learn/iaas-paas-saas

21. Kannan P., (2015). Beyond hadoopma preduce apachetez and apache spark. San Jose State University. http://www.sjsu.edu/people/robert.chun/courses/CS259Fall2013/s3/F.pdf

22. Li. Xue, G. Liu, J. Parfitt, X. Liu, E. Van Herpen, Å. Stenmarck, C. O'Connor, K. Östergren, S. Cheng,.(2017). Missing food, missing data? A critical review of global food losses and food waste data, Environ. Sci. Technol., 51 (2017), 6618-6633, https://doi.org/10.1021/acs.est.7b00401.

23. Majumdar, J., Naraseeyappa, S. & Ankalaki, S. (2017). Analysis of agriculture data using data mining techniques: application of big data. J Big Data. 4, 20. https://doi.org/10.1186/s40537-017-0077-4.

24. Mantle Labs, https://www.mantle-labs.com

25. McCue, L. (2020, December 16). Supply chain analytics: What it is and why it matters. Oracle NetSuite. https://www.netsuite.com/portal/resource/articles/erp/supply-chain-analytics.shtml

26. Mukherjee Sanjeeb. (Tuesday, January 17, 2023). PM Narendra Modi withdraws three farm laws, asks farmers to go home. *Business Standard*. https://www.business-standard.com/article/current-affairs/pm-narendra-modi-withdraws-three-farm-laws-asks-farmers-to-go-home-121111900296_1.html

27. Pantazi, X.E., et al. (2016). Wheat Yield Prediction Using Machine Learning and Advanced Sensing Techniques. Computers and electronics in agriculture, V.121, 57-65. https://doi.org/10.1016/j.compag.2015.11.018

28. Project FarmVibes. Democratizing digital tools for sustainable agriculture, https://www.microsoft.com/en-us/research/project/project-farmvibes/

29. ScienceDaily, (10 January 2023). Project aims to expand language technologies: Research could bring automatic speech recognition to 2,000. Carnegie Mellon University. https://www.science-daily.com/releases/2023/01/230110151049.htm

30. Seema Maitrey, & C.K. Jha. (2015). MapReduce: Simplified Data Analysis of Big Data, Procedia Computer Science, Volume 57, 563-571, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2015.07.392.

31. Spark Core Programming. https://www.tutorialspoint.com/apache_spark/apache_spark_rdd.htm

32. Srivastava, A., (2018). Technology Assisted Knowledge Agriculture for Sustainable Development Goals. Adv Crop Sci Tech 6, 391. https://doi.org/10.4172/2329-8863.1000391
33. Swarna, C. & Ansari, Zahid. (2017). Apache Pig - A Data Flow Framework Based on Hadoop Map Reduce. International Journal of Engineering Trends and Technology. 50. 271-275. https://doi.org/10.14445/22315381/IJETT-V50P244
34. Tableau, https://www.tableau.com
35. Talend, Big Data and Agriculture: A Complete Guide, https://www.talend.com/products/data-fabric/
36. Taranis, https://taranis.ag
37. TensorFlow, https://www.tensorflow.org
38. The Economic Time. (Dec 24, 2022). Govt to provide free food to 814 m Indians till Dec '23. https://economictimes.indiatimes.com/news/india/govt-to-provide-free-food-to-814-m-indians-till-dec-23/articleshow/96468364.cms?utm_source=contentofinterest&utm_medium=text&utm_campaign=cppst
39. The History of Apache HBase: A Complete Guide, https://history-computer.com/the-history-of-apache-hbase-a-complete-guide/
40. U.S. Environmental Protection Agency. Population Reference Bureau (PRB). (2016). World Population Data Sheet. https://www.prb.org/wp-content/uploads/2016/08/prb-wpds2016.pdf
41. Veenadhari, S., Misra, B. and Singh C., (2014). Machine learning approach for forecasting crop yield based on climatic parameters, International Conference on Computer Communication and Informatics, Coimbatore, India, 2014,1-5, https://doi.org/10.1109/ICCCI.2014.6921718
42. WDF University for SDG2030 https://wdfuniversity.org
43. Wilde Matt. (2021).Progressive Farmer Crops Editor. Farmland Loss Threatens Food Supply. https://www.dtnpf.com/agriculture/web/ag/news/article/2021/02/01/land-use-farmland-loss-threatens
44. WWF report. (21 July 2021). Over 15% of food is lost before leaving the farm. https://www.wwf.eu/?4049841/fifteen-per-cent-of-food-is-lost-before-leaving-the-farm-WWF-report