Global Journal of Computer Science and Technology
Neural & Artificial Intelligence
Volume 12 Issue 10 Version 1.0 Year 2012
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

# Primary Education Status Analysis in Bangladesh Based on Neural Networks and Baysian Networks

By Snehasish Sarker, Md.Sarwar Kamal & Puja Das
*BGC Trust University*

*Abstract -* In this research work we have concentrate to measure the primary education status in Bangladesh, a developing country of South Asia. It known that the literacy rate of South Asian country is very slow and it is not the different in Bangladesh. Here we measure the dropout rate of primary school kids at different classes at different sessions. We have collected the data from various primary schools from Chittagong region of Bangladesh. Here we use K –Nearest Neighbor (KNN) algorithm to classify the data from irrelevant data like secondary school and tertiary level data. After then we have applied Neural Network (NN) to train the data set for better result. Finally we have compared the result by calculating the result with Bayesian Network (BN). Here we found that if the dropout rate is small Neural Network is best to measure the result and NN generate more error when the dropout rate is large. On the contrary BN is better when the rate is large.

*Keywords :* *K-Nearest Neighbor Algorithm, Neural Network, Bayesian Network (BN), Primary school, Dropout rate.*

*GJCST-E Classification:* *C.1.3*


*Strictly as per the compliance and regulations of:*

# Primary Education Status Analysis in Bangladesh Based on Neural Networks and Baysian Networks

Snehasish Sarker [α], Md.Sarwar Kamal [σ] & Puja Das [ρ]

*Abstract* - In this research work we have concentrate to measure the primary education status in Bangladesh, a developing country of South Asia. It known that the literacy rate of South Asian country is very slow and it is not the different in Bangladesh. Here we measure the dropout rate of primary school kids at different classes at different sessions. We have collected the data from various primary schools from Chittagong region of Bangladesh. Here we use K –Nearest Neighbor (KNN) algorithm to classify the data from irrelevant data like secondary school and tertiary level data. After then we have applied Neural Network (NN) to train the data set for better result. Finally we have compared the result by calculating the result with Bayesian Network (BN). Here we found that if the dropout rate is small Neural Network is best to measure the result and NN generate more error when the dropout rate is large. On the contrary BN is better when the rate is large.

*Keywords : K-Nearest Neighbor Algorithm, Neural Network, Bayesian Network (BN), Primary school, Dropout rate.*

## I. Introduction

The development of any country depend s mainly on its manpower and the pillar of good manpower is the primary education. It is easy to realize that as the children are engaging to primary education as the hopes of the prosperity will go up for respective country. Kids learning are very important for every country irrespective of rich and poor. The countries which are more developed are developed at their education level and the development start from primary level. Basically kids are very much curious in every matter and they loves to adopt with innovative culture and fashion. To ensure the kids literacy governments as well as the parents should take effective measure which will attract the kids to learn with joy and enjoyments.

Proper education for the kids can empower human beings to liberate individual mind from the curse of ignorance and darkness. It represents the foundation in the development process of any society and the key indicator of the people's progress and prosperity.

*Author α ρ : Student of computer science and engineering at BGC Trust university, ID: 0913006, 0913009.*
*Author σ : Lecturer, computer science and engineering, BGC Trust University.*
*E-mail α : snehasish.bgc@gmail.com*
*E-mail σ : puja.cse.bgc@gmail.com*
*E-mail ρ : sarwar.saubdcoxbazar@gmail.com*

In the view of the importance of education to a country like Bangladesh the present thesis addresses limitations of primary education system, which is diversified and multifarious due to economic, socio cultural, political, regional and religious factors. The entrance of primary education is maintained mainly by the government of Bangladesh. More than 75% schools are controlled by the government and around 83% of the total children enrolled in the primary level educational institution go to these schools (Baseline Survey, 2005:3). Similarly, more than 70% primary teachers are working in the government controlled schools. Besides government run primary schools, nine other category of primary schools are administered, monitored and maintained by different authorities. Disparity and lack of coordination among these institutions constrains the attainment of universal primary education and in its effort to increase enrollments and quality education. Variations in teacher student ratio, the number of qualified and trained teachers between the categories, also pose a big challenge towards achieving the goals of universal primary education.

In the circumstances of the open scenario, Bangladesh became one of the signatories to the UN Millennium Declaration in 2000, and has achieved to eight Millennium Development Goals (MGDs) that affirmed a perception for the 21st century (Burns et al, 2003:23). Bangladesh also pledged to implement the MDGs roadmap by 2015. The MDG-2 targets for 'Achieving Universal Primary Education' are claimed to be on track in Bangladesh, showing remarkable achievements in terms of net enrolment rate in primary education 73.7% in 1992 to 87% in 2005 and primary education completion 42.5% in 1992 to 83.3% in 2004 (Titumir, 2005:120). Bangladesh government itself had taken many initiatives, including the Compulsory Primary Education Act 1993, which made the five-year primary education program free in all primary school. The government adopted demand side intervention policies such as food for education program and stipend program for primary education. Of late, the government introduced primary education development program (PEDP-II), a six year program beginning in the year 2000, which aims to increase access, quality and efficiency across the board in the primary education sector. Despite existing socioeconomic problems,

Bangladesh by now has achieved a good progress in net enrollment rate and education completion rate in primary education. The current paper will examine the outcomes and challenges that have emerged as a result of Bangladesh being the signatory of MDG-2 for achieving universal primary education by 2015. It would further investigate whether the target of the second Millennium Development Goal is attainable within the stipulated time.

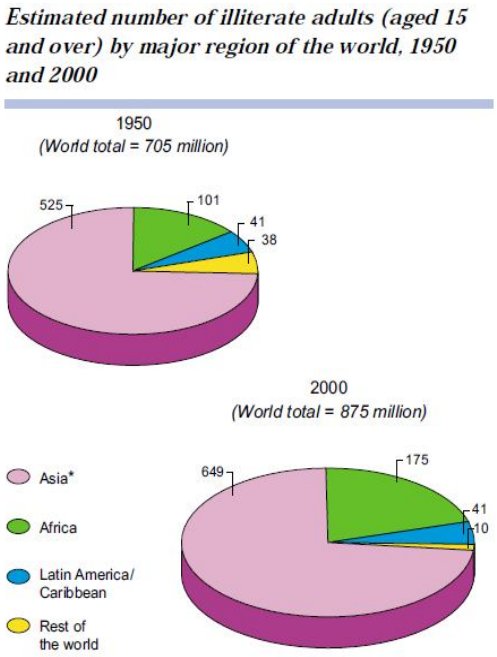
To better improvement of literacy at primary level we examined and interviewed a lot of kids at various places and found that the children are used to learn with joy and entrainment. Recently in Bangladesh various primary school at North Bengal part of Bangladesh implemented park at primary school compartment and get the very positive results. Hathazari Upazila parishad has arranged an innovative approach to help bring down dropouts in government primary schools. It is setting up children's parks in schools and constructing roads to give students easy access to schools. Encouraged by cent percent pass rate in last year's class-V terminal examinations in the upazila in Patuakhali, the parishad set up children's parks in 60 of 122 primary schools [3] there to draw students. Construction of 50 more is underway. It also paved 25 roads for easy access to schools for [3] students from nearby areas. About 65 more are under construction. Urmee, a student of class-IV at Madanpura Government Primary School said she could not attend school in the rainy season last year for bad condition of the road to her school. But she has not missed any classes this year after the road got repaired. Miss Dhunu Chokrabarty headmaster of the Mujaffarabad govt primary school said most students now come to school in the rainy season but the attendance hovered around 70-75 percent during the same period before the road was repaired. The children's park in the school is an additional attraction for students, he added. Several students were found waiting for their turn to get on a swing at Hosnabad Government Primary School. Many of them come to school early to play on the swings. Rafiqul Islam, a student of class-IV of Najirpur Government Primary School, said their playground used to go under water in the rainy season. It was developed last year, and now they can use it all year round Dilip Chandra Sarker, headmaster of Baromashia govt primary school, said the parish ad's initiative helped increase attendance rate in his school. Now about 93 to 97 percent students attend classes, while it was 80 to 85 percent a year ago. Ayesha Akter Chowdhury,[3] headmaster of Sholkata govt primary school, said the children's park of the school has brought a big change. The attendance rate went up to more than 90 percent from below 80 percent. Upazila Chairman Engineer Mojibur Rahman said they are setting up children's park on school playgrounds to provide students with leisure facilities on instructions from Patuakhali-2 lawmaker ASM Feroz, also whip of the Jatiya Sangsad. The parishad is spending about Tk 50,000 on each children park from its own fund. The project started in January 2010 and will end by December next year."We have already re-excavated 36 school ponds. Of them, 10 have been brought under fish cultivation with fisheries department's help. Another 63 will be re-excavated by next year," said Mojibur, who last year received an award from the prime minister for his role in promoting primary education in his area. "We will take another programme to grow seasonal vegetables on unutilised school land. The profits from the sale would go to poor and helpless students for buying books and stationery," he said. The parishad published a 472-page book Tathya Kanika. It contains names of all educational institutions in the Upazila and information on students and teachers."We have already distributed the book to all educational institutions in the area. It will help students get in touch with each other. Upazila officials will also be able to communicate with any teachers or institutions by using the information provided in the book," said Mojibur.Md Ibrahim, Upazila primary education officer, said "We are getting good results from these steps."Now most students go to school that has become a more interesting place for them, he added.

The organizations of the work start by the literature overview after the introductory descriptions. At the introductory description we have observed the situation of Bangladesh primary education and the current status of the country. At the literature study part we have look forward towards the whole world situations specially the developing courtiers of the world like Ethiopia, Sudan, Nigeria and Nepal. The UNESCO report also helps us a lot towards the exact scenario of the education status of the developing country. The data collection is done after the literature study and the data set are real world data from various primary school of Chittagong region of Bangladesh. It is very alarming that we have found very much irregularity to collect the data. Data collection helps us to design the intelligent system for our desired work. At first we applied the data classification techniques to the collected data. We choose K-Nearest Neighbors (KNN) algorithm for data set classification. The data set classification helps us to reduce the redundant data. At last but not the least we compare the result by both Neural Networks (NN) and Bayesian Networks (BN).

## II. Literature Over View

We have studied and checked related supporting documents and information towards the education status analysis throughout the world irrespective of rich and poor country. But in this work we have concentrate to design and investigate the condition of developing country. We found that there are some paper support the qualitative measurements and

some other had concentrate on survey based measurements. From the study of Ethiopia, one of the backward countries at education and economic condition we see that only survey based analysis is done [1]. Data are collected from students, instructors, gender officer and from guardian though the interview, questionnaires and discussion with focus group. Here no advanced technique is used like Neural Network (NN) and Bayesian Network (BN) or data mining techniques as data pre-processing, data warehousing, linear regression, non linear regression and so on. But for the generic evaluation and measurement it is very essential to have to have an Intelligent Decision System (IDS) at every levels of education. According to the World Education Report 2000 we have noticed that at South Asia there is huge part of the children are out of education and they are suffering a lot of socio-economic problems like poverty, proliferating acts of violence and conflicts, illiteracy and rich and poor gap. The survey by the expert of UNESCO found and 649 [2] million illiterate resides at south Asian region. The figure 1 bellow depicts the matter.




* For 1950, data for the Asian countries of the former USSR are included in Rest of the world.
Source: Figures for 1950 are taken from World Illiteracy at Mid-Century, p. 15, Paris, UNESCO, 1957.

*Fig. 1 :* The illiteracy comparison at 1950 and 2000

All the reports of UNESCO and survey paper of the researcher are based on people interaction and qualitative. But run the quick and accurate result it is convenient to have an intelligent system. If we want to perform the calculation to the Robot what will be situation than? We must need the Decision Support System (DSS).

## III. Data Collection

We have collected the data set from the Chittagong region of Bangladesh for the experiment of our research work which is done by the assistance of continues support by the head teacher of the schools. We mainly concentrate data set for the classes where the students enrolled at first class and those are appeared for the Primary School Certificate (P.S.C) exam. In this condition we found that the measurement is efficient and meaningful for the evaluation as well as for the decision. Our sample spaces were the Mujaffarabad Government Primary School, Hathazari, Baromashia Government Primary School, Fatikchari, Sholkata Government Primary School, and Anwara. From Mujaffarabad Government Primary School, Hathazari we have collected consecutive data from 2004 to 2008 for each year from class one to class five. The table 1 to 5 narrates the collected data for various batches from 2004 to 2012. For table 1 we collect the data set for class one to class five for a respective batch. Here we can see that from 2004 to 2005 four students have dropout among thirty students. At 2006 we see that there are more students than class one due to the dropout of the previous class those who are enrolled at class one at 2003. For a certain batch of 2004 at class one. We get some variations at class three, four and five due to the reasons of dropout of the previous classes.

| Year | Class | Total Students |
|---|---|---|
| 2004 | Class 1 | 30 |
| 2005 | Class 2 | 26 |
| 2006 | Class 3 | 35 |
| 2007 | Class 4 | 32 |
| 2008 | Class 5 | 33 |

*Table 1 :* Mujaffarabad Government Primary School, Hathazari

As the same table 2 depicts the data set for the same school for another batch that is started from 2005 for class one and end the class five at 2009.

| Year | Class | Total Students |
|---|---|---|
| 2005 | Class 1 | 37 |
| 2006 | Class 2 | 33 |
| 2007 | Class 3 | 40 |
| 2008 | Class 4 | 46 |
| 2009 | Class 5 | 36 |

*Table 2 :* Mujaffarabad Government Primary School, Hathazari

This table also indicates the dropout rate from class three to class five and some new students also get enrolled from class three to class five. It is really pathetic to see that at village part the kids are out of school at various ages due to the lack of information or other socio-economic problems. The both tables above indicate the data set from rural part of the Bangladesh.

The table 3 below also shows the same data set for the same school but the timeline is different than other two tables. Here we see that the enrollment rate increase over the time. It is also positive side for Bangladesh that the students and the guardians are now becoming causes than earlier.

| Year | Class | Total Students |
|---|---|---|
| 2006 | Class 1 | 45 |
| 2007 | Class 2 | 38 |
| 2008 | Class 3 | 32 |
| 2009 | Class 4 | 29 |
| 2010 | Class 5 | 26 |

*Table 3 :* Mujaffarabad Government Primary School, Hathazari

We than collect the data set from the area where the people are more causes than the previous area and we have observed a interesting change towards the developments of the literacy. At this place the government involvements are also very frequent than the previous area.

| Year | Class | Total Students |
|---|---|---|
| 2007 | Class 1 | 129 |
| 2008 | Class 2 | 119 |
| 2009 | Class 3 | 112 |
| 2010 | Class 4 | 98 |
| 2011 | Class 5 | 85 |

*Table 4 :* Baromashia Government Primary School, Fatikchari

The table 4 contains the data more than other three tables and the dropout is also more than the others. If the data size is more the dropout is also more. The main reasons behind the dropout is that the poverty and guardians inconsistency for study and literacy. Besides they have the idea what will happen after study, for job they needs lobbing. Here for Fatikchari area we have found that majority of the people are living at middle east for earning as result the students enrollment rate increase for the better financial support. Similarly at table 5 the enrollment is smaller than the Fatikchari area due to the same facts.

| Year | Class | Total Students |
|---|---|---|
| 2006 | Class 1 | 65 |
| 2007 | Class 2 | 57 |
| 2008 | Class 3 | 50 |
| 2009 | Class 4 | 48 |
| 2010 | Class 5 | 38 |

*Table 5 :* Sholkata Government Primary School, Anwara

## IV. K-Nearest Neighbor (K-NN)

K–nearest neighbor (K-nn) algorithm is a branch of supervised learning. Now-a-days it is being applying in various fields of data and information processing irrespective of science, commerce and arts. In the context of machine learning, K-nn is considered an effective data classification technique based on adjacent developed examples of sample space. The value of K is always positive and an object is classified by considering the greater number of choice of its neighbors. The neighbors are chosen from data set which is best fit for correct classifications and Euclidean distance helps to measure the overall distances. Here every occurrence correlates to points in sample space or within populations. Generally distance or similarity between instances or objects is easy if the data sets are numeric or integer. A very typical formula to calculate distances is Euclidian distances formula as follows:

$$d = \sqrt{((x_{1i}-x_{1j})2+ (x_{2i}-x_{2j})2)} \quad (a)$$

In some cases Manhattan or City Block distance also applicable:

$$d = (x1i-x1j) + (x2i-x2j) \quad (b)$$

However it is very essential to bear in mind that all the instances at sample space must be same scale. As for example income will compare with income not the height of the human beings.

For qualitative data the distance measurement process will be different and it is important to consider that the instances are same or not. At this stage the qualitative objects are measured by allocating Boolean values to each object. It might be possible to converts to instances between which distance can be identified by some techniques. As for example color, temperature, age, height etc. Text and character has identified as one instance per word with the frequency start from 0, 1, 2……n.

*a) The classifications process of K-nn as follows*

The two main steps of K-nn must follow are:

1. Training
2. Predictions

Training means to get information from all sample spaces and populations. To accomplish this work we need to have the idea about the all instances and objects. In this sense it is very important to bear in mind that data set must be in same class. The qualitative and quantitative data measurement will be different.

The predictions will manage by considering the predefined methods.

*b) The k-nn Algorithm*

The total algorithmic steps are as follows:

1. Parameter selections (int m, int n). m=0, n=1, 2, 3……n.
2. Distance calculation

   $\sqrt{\Sigma}$ (q-pi) 2 where i=0,1,2,3……..,n

3. Short the distances of sample space and marked the closet neighbors in the context of K-th smallest distance.

SHORT NEIGHBORS (**S**, **C**)
**Input** instances **S** with n sample objects, comparator **C**
**Output** instances S sorted according to **C**
**if** S.length() > 1
Then ($S_1$, $S_2$)←**divide**(S, n/2)
SHORT NEIGHBORS ($S_1$, C)
SHORT NEIGHBORS ($S_2$, C)
S ← SHORT NEIGHBORS ($S_1$, $S_2$)

4. Similarities assumption :Instances that are close together should have similar values Minimize

$$\xi(f) = \sum w_{ij}(f_i - f_j)^2$$

Where $w_{ij}$ is the similarity between examples i and j.
And $f_i$ and $f_j$ are the predictions for example i and j.

5. Predict the value as follows:

Standard KNN $\hat{y} = \arg\max_y C(y, Neighbors(x))$

$$C(y, D') \equiv |\{(x', y') \in D' : y' = y\}|$$

Distance-weighted KNN

$$\hat{y} = \arg\max_y C(y, Neighbors(x))$$

$$C(y, D') \equiv \sum_{\{(x', y') \in D' : y' = y\}} (SIM(x, x'))$$

$$SIM(x, x') \equiv 1 - \Delta(x, x')$$

6. Find out the best heuristics distance

```
f(n) = g(n) + h(n)
```

Where:

- g(n) is the cost of the best path found so far to n
- h(n) is an admissible heuristic
- f(n) is the estimated cost of cheapest solution through n

## V. New Maximum Nearest Area (NMNA)

How k-nn selects the desired values from a lot of alternatives is that it calculates its nearest most predicted value. The following figure depicts the computations.

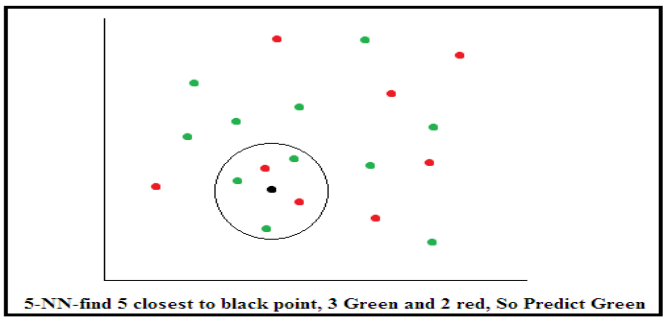



*Fig. 2 :* The neighbor selecting process

In the figure above we see that the small circle belongs three different color dots where the black one is the pivotal element and based on that point we will calculate the green and other two green and red points. According to this figure we have to predict the green points as a K nearest neighbors. The neighbors are very closest to the pivotal point. It is vital point that New Maximum Nearest Area (NMNA) must be accurate otherwise it will not work properly for data set selections. In the k-nearest neighbour process, the only benefit of selecting a large k value is to scale down the variance of the conditional probability estimate. By parallel, in oursystem, a large k value can probably lead to a large confidence measure, therefore to a small probability of error. It is very much essential for each inquiry point of the confidence measure and the corresponding probability of error can be easily computed for different numbers of neighbors. Therefore, one can choose to increase the neighborhood size k until a preset probability of error threshold is achieved. So, the probability of error, or equivalently the confidence level, provides a mechanism to dynamically determine the size of the neighborhood. We will call the modified version of the k-nearest-neighbor rule the confident-nearest neighbor rule.

## VI. Organization of the Process

Now it is important to build the process how K-nn may organized in reality or the time line. To manage the proper training area we have to shorten the area or to select the appropriate area. When we are able to fix the sample area for computation, it will help us to reduce the computational complexity for entire process. To accomplish the total work we must follow some preconditions. First we have to choose a population area where we can apply our algorithm to extract our outcome. The figure 2 bellow shows a population area for our research activity.

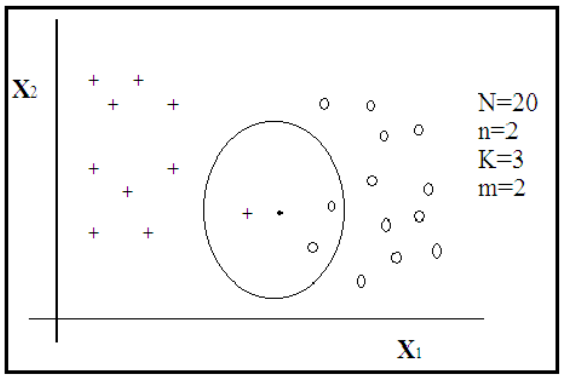



*Fig. 3 :* The total population area for K nearest neighbors

Here,
N = Total number of data set at population space. In this figure above we see that there is twenty (20) objects are outside the circle. The circle denotes the selected sample space. Inside the circle the black point indicate the pivotal or central point.

K=the total neighbors. Here the value of K is three (3).
n= indicate the nearest value.
m= categories of the neighbors. In the figure above
We see that there are two categories of neighbors.
One data set indicate by plus (+) sign and other is small hole.

At the beginning we narrow the area as sample space from population area. The figure 3 below shows the desired sample space for our working activity. By the reduction process, we are able to reduce forty percent of the computational cost.

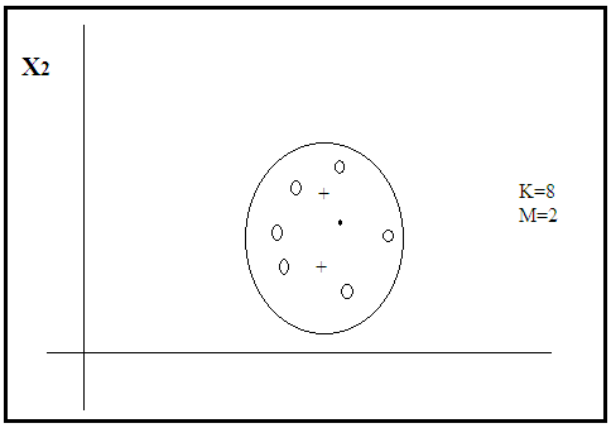



*Fig. 4 :* The reduction sample area for K nearest neighbors area

By comparing both figure 2 and figure 3 we get the followings equation as a general format.

$$\text{Computational reduction cost (CRC)} = \frac{Samplespaceneighbors}{Populationneighbors}.$$

Suppose at figure 2 we see that there are twenty objects or data set and after making classification based on K nearest neighbors algorithmic techniques we have extracted as figure 2 where there is only eight data set or object which are very much closed to pivot point indicated inside the circle as black point.

$$\text{Here the CRC} = \frac{08}{20} \times 100$$

$$= 40\ \%$$

It is really interesting and effective that for twenty data points we have reduced 40% of the computational cost and as a consequence it reduced the complexity of the computation. From the time and space complexity view point in linear search we get the better improvement at both time and space complexity. Time complexity defines how the algorithm behaves while the input size increases. Generally in the worst case, the running time is proportional to n where n denotes as input size. For any input size with time complexity O(n) will take the twice time while the input size doubles as a consequence with time complexity O(N2) will take four times longer while input doubles and so on. As a average case the time complexity is as follows:

C (n) =1*1/n+2*1/n+3*1/n+.....+n*1/n = (n+1)/2.
So here we get the improvements of (n+1)/2 * 40 %
=(n+1)/2.*40/100
=40* (n+1)/200.
= (n+1)/5.

Hence we see that the CRC helps to cut the cost to one fifth of the original cost. So the time complexity is the CRC linear search is O (n/5) where the total population cost is O (n) for the n input size. It will be easy to remember that the result of comparison will change when the input size will change. From further reduction we get the figure below at figure 4. According to the equation a we manage our desired sample data set configured as bellow.

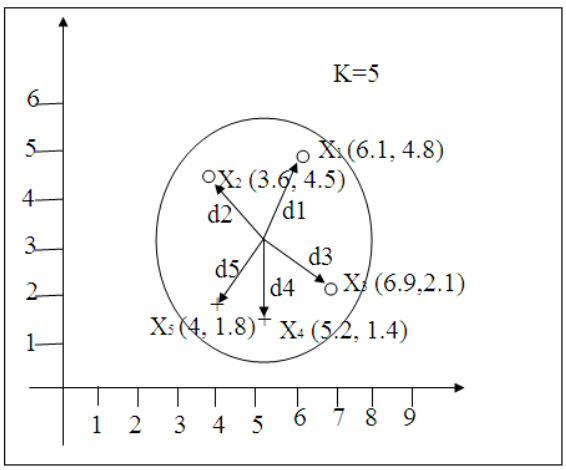



*Figure 5 :* The final reduction of the K nearest neighbors

At figure five we depict the problem domain by a time line diagram. The time line diagram behaves the similarities of Support Vector Machines (SVM) of Maximum Margin Hyper-plane (MMH). The pivot value is selected as the mean point of the data set.

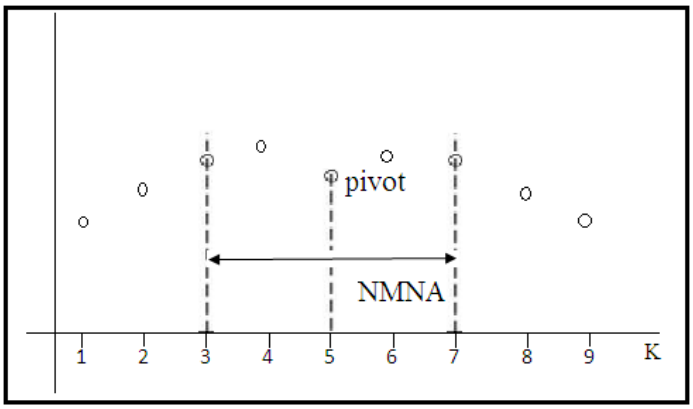



*Figure 6 :* The time diagram of K nearest neighbours as a MMH of SVM

## VII. Neural Networks Perception

Neural networks represent a brain analogy for information processing. These models are biologically exhilarated rather than clear-cut clone of how the brain actually functions. The figure 7 shows the similarities between artificial neural network and biological neurons. Neural ideas are usually implemented as system simulations of the massively parallel processes that involve processing elements interconnected in network architecture.

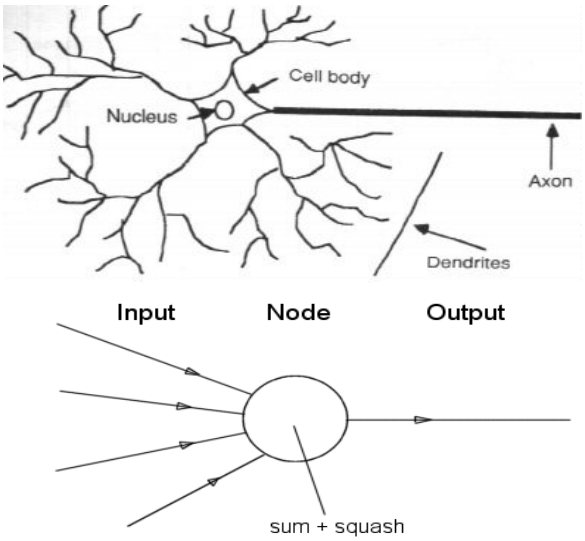



*Fig. 7 :* The biological and artificial neurons

Neurons receive the sum of information from other neurons or the external elements, perform transformation on the inputs and then pass the transformed information to other neurons or the external outputs. A typical structure is shown in Figure 8. For the better measurement and accurate result we have experimented by Multi Layer Involvement (MLI) of the Neural Networks (NNs) which is an advanced computational and learning method at modern computation and Intelligent Systems (ISs). A MLI consists of three layers named input layer, hidden layer, and output layer. A hidden layer receives input from the previous layer and converts those inputs into outputs for further processing. Several hidden layers can be placed between the input and output layers, although it is quite common to use only one hidden layer. Every working cell is connected with each of other cell as directed graph. A NN is very much similar with a directed graph where the neuron cells are considered as vertices and the connections between the cells are edges. In that case each edge is associated with its weights and the weights must reflect the measurements of the input and output results. Naturally an Artificial Neural Network (ANN) is consisting of Adaptive Linear Neural Elements (ADLINE) that changes its structure according to the propagation of information on external or internal matters through the network during the learning phase.

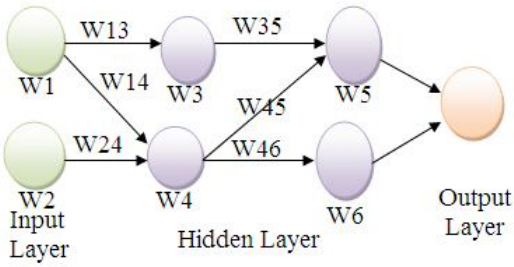



*Fig. 8 :* Multi layer neural network

Neural Networks learn to adapt the inputs to produce desired outputs. When the NNs choose to learn by supervised process than the learning must be inductive. To learn the NN first compute the temporary output, then compare the output with target output and finally adjust the weight and repeat the process. When existing output are available for comparison, the NN start the learning process. The figure 9 bellow shows the learning process of a single neuron. The weight adjustment with learning is $\Delta W_i = \eta * (D-Y).I_i$.

Where,

$\eta$=learning rate, D= Expected output, Y= actual output, I= input.

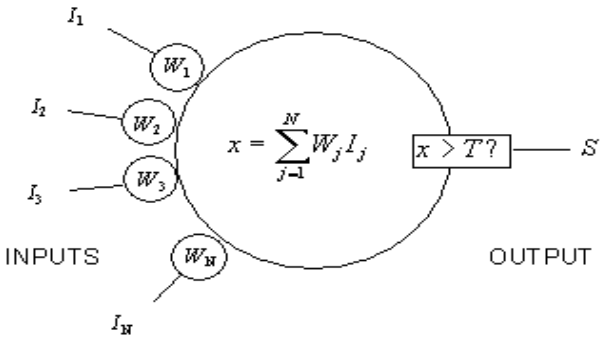



*Fig. 9 :* The learning process for a single artificial neuron

For a Multi Layer Neural networks the mathematical learning process is as follows:

$$y_k^1 = \frac{1}{1+e^{-w^{1kT}x-a_k^1}}, k = 1,2,3 \quad \text{(c)}$$

$$y^1 = (y_1^1, y_2^1, y_3^1)^T$$

$$y_k^2 = \frac{1}{1+e^{-w^{2kT}y^1-a_k^2}}, k = 1,2$$

$$y^2 = (y_1^2, y_2^2)^T$$

$$y_{out} = \sum_{k=1}^{2} w_k^3 y_k^2 = w^{3T} y^2 \quad \text{(d)}$$

For error measurements the equation is

$$E = \frac{1}{N}\sum_{t=1}^{N}(F(x_t;W) - y_t)^2 \qquad (e)$$

To change the weight the learning equation are

$$\Delta w_i^j = -c \cdot \frac{\partial E}{\partial w_i^j}(W) \qquad (f)$$

$$w_i^{j,new} = w_i^j + \Delta w_i^j$$

C is a learning parameter. Usually it is constant.

## VIII. Bayes'Theorem

Bayes' theorem and conditional probability are opposite to each other. Given two dependent events A and B. The conditional probability of P (A and B) or P (B/A) will be P (A and B)/P (A). Related to this formula a rule is developed by the English Presbyterian minister Thomas Bayes (1702-61).According to the Bayes rule it is possible to determine the various probabilities of the first event given the outcome of the second event in a sequence of two events.

The conditional probability:

$$\text{P(B/A)} = \frac{P(AandB)}{P(A)} \qquad (1)$$

The equation (1) will help to find out the probabilities of B after being occurrences of the A. we get the Bayes' theorem for these two events as follows:

$$\text{P (A/B)} = \frac{P(A).P(B/A)}{P(B)} \qquad (2)$$

If there are more events like A1, A2, and B1, B2.In this case the Bayes theorem to determine the probability of A1 based on B1will be as follows:

P(A1/B1)=

$$\frac{P(A1).P(B1/A1)}{P(A1).P(B1/A1) + P(A2).P(B2/A2)}$$

## IX. Implementation

The implantation is done according to the concepts and knowledge of NNs and Bayesian Networks. The following figure 11 narrates the brief descriptions of the implementation. The fundamental algorithmic steps perform as the core idea or engine for the calculation. Our proposed algorithm for dropout prediction is as follows:

### The algorithmic steps (Proposed)

1. Start
2. Take node number n.
3. Find out the probability P(A) for all nodes.
   P(A)=((A-b)/A)*100;
4. Find out the average probability for all nodes.
   Average probability = (P(A)+P(B)+…….+P(N))/n;
5. Let the average probability is equal to the probability of last node. P(Y)= Average probability of all nodes
6. Find out the result node X.
   X=Y-((P(Y)*Y)/100);
7. End
8. Repeat steps 2 to steps 6 again.

To accomplish the full work with efficiency and accurate we have worked firstly on input data sets and fit them to the networks. In this work we set input as a numeric data for the Neural Network. Here four inputs are applied to the network to predict the output of the class five. Our processing shows in figure 11.

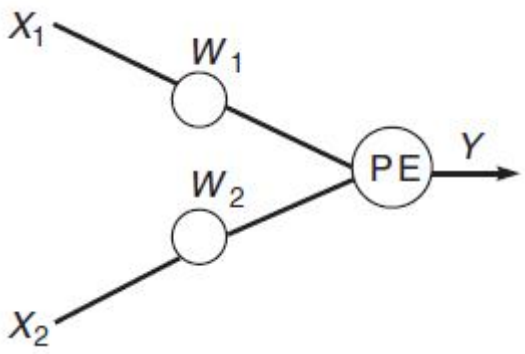



*Fig. 10 :* A processing element with two inputs set

X1 and X2 may be the numeric value or representations of an attribute. W1 and W2 are the weights for the networks and Y denotes the output of the inputs. PE indicates the Processing Element. The simple calculation for the output Y = X1 W 1 + X2 W 2. The output indicates the salutation of the inputs. At the output level the ANN provides the value 1 for yes and 0 for no. another important part for the processing is weights of the network. Weights are the relative strength of the inputs set or many weights transfer input from layer to layer.

```
$w=((($a-$b)/$a)*100);
$x=((($b-$c)/$b)*100);
$y=((($c-$d)/$c)*100);
$z=(($w+$x+$y)/3);
$e=((($z*$d)/100)+$d);
echo "According to Newral network the result is ".$e;
```

*Fig. 11 :* The brief idea for NN implementation

One of the most important functionality of this work is the error learning training to get the desired outputs. The figure 12 shows the first pass of our research work. Gradient of the neuron= **G** =slope of the transfer function×[Σ{(weight of the neuron to the next neuron) × (output of the neuron)}].

G1=(0.7265)(1-0.7265)(0.0397)(0.5)(2)=0.0093.
G2=(0.6508)(1-0.6508)(0.3492)(0.5)=0.0397.
G3=(1)(0.4292)=0.4292. Error =1-0.4292=0.5708.

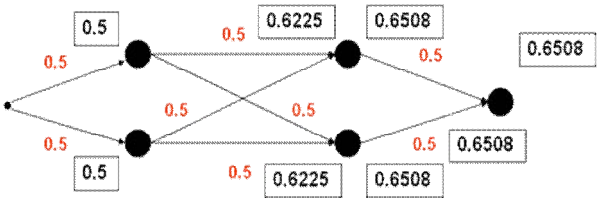



*Fig. 12 :* Computation at NN for the collected data set

To minimize the errors rate we have adjusted the weights which have shown bellow at figure 13.the new weights is New Weight=Old Weight + {(learning rate)(gradient)(prior output)}.
0.5+(0.5)(0.0093)(1),0.5+(0.5)(0.0397)(0.6225,
0.5+(0.5)(0.3492)(0.6508).

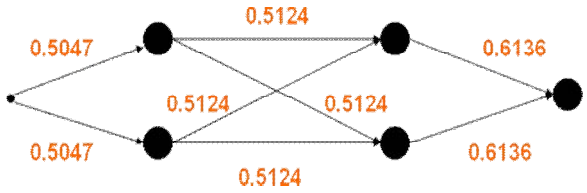



*Figure 13 :* The weight adjustments

Finally we have implemented for Bayesian network and we founds as figure 14. According to the operation of the Bayesian networks and application in this work we have noticed that BN runs as follows.

```
$total=($a+$b+$c+$d);
$i=($b/$total);
$j=($c/$total);
$k=($d/$total);
$l=(($i+$j+$k)/3);
$r=($l*$total);
echo "According to Baye's theuram the result is ".$r;
```

*Figure 14 :* The Bayesian implementations

## X. Result

The results of this implementation snaps are shows bellow.

Input students number in the text field

Class 1: 129
Class 2: 119
Class 3: 112
Class 4: 98
Submit

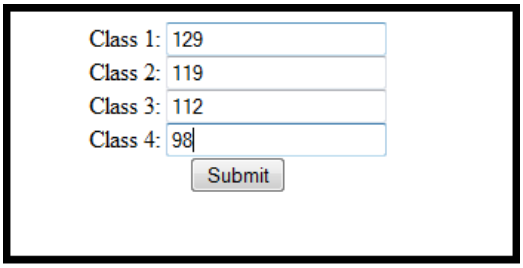



*Figure 15 (a) :* The inputs for the Baromashia School

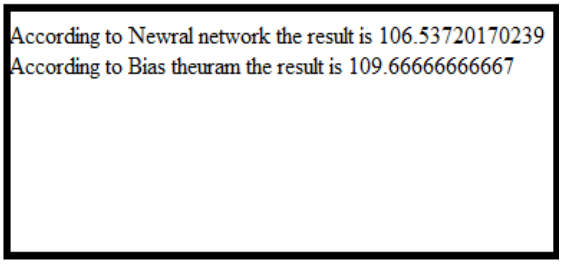



*Figure 15(b) :* The outputs of the previous inputs

Input students number in the text field

Class 1: 30
Class 2: 26
Class 3: 35
Class 4: 32
Submit

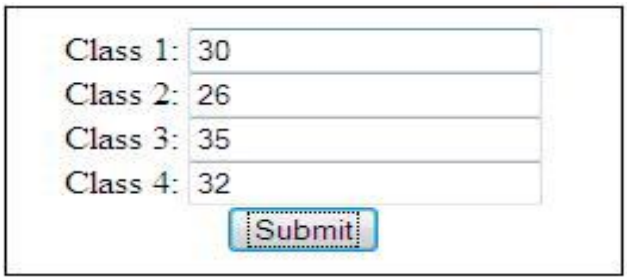



*Figure 16 (a) :* Inputs for Mujaffarabad School

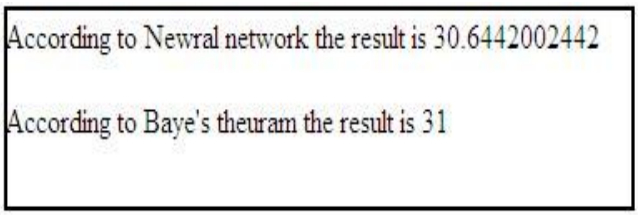



*Figure 16 (b) :* The outputs of the previous inputs

## XI. Comparisons

The result of the both NN and BN are very effective for this research. Both procedure are very suitable for machine to learn and predicts the accuracy of the result. Machine learning is very essential for current era. In the age of information superhighway every steps of computations are becoming machine oriented and are being handing based on automated way. Though the efficiency we make a comparison and found that NN is better at the cases when the input size has the less difference in dropout rate and on the contrary the BN is better while the input dropout rate is larger.

NN is 96% is corrects and BN is 91.5% accurate in this research.

## XII. Acknowlefgement

We are thanking the head masters of the schools who help us by providing the data regarding the enrollments and dropout students from their schools. We are showing our respect to Miss Dhunu Chokrabarty sir, head teacher of Mujaffarabad govt primary school, Mr. Dilip Chandra Sarker sir head teacher of Baromashia govt primary school, and Miss. Ayesha Akter Chowdhury head mistress of Sholkata govt primary school for their regular support for data collection and analysis.

## XIII. Conclusion

It is very good to say that this research will help to assess the degree of dropout students from any country especially from the developing country. We have measured the significant amount of dropout students from primary stage due to the various socioeconomic problems like lack of knowledge, poverty, and social barrier. Our implementation is very efficient for automated system as well as machine learning. Besides this work we noticed a few drawback regarding the time line and data collections and organization. At future we will overcome the problems regarding the indicated problems.

## References Références Referencias

1. Trend and causes of female student's dropout from teacher education Institutions of Ethiopia: the case of Jimma university by Wudu Melese, Getahun Fenta. Ethiop. J. Educ. & Sc. Vol. 5 No. 1 September, 2009.
2. World education report 2000 by UNESCO.
3. The daily star Bangladesh.
4. NEURAL NETWORKS IN DATA MINING by DR. YASHPAL SINGH, ALOK SINGH CHAUHAN, Journal of Theoretical and Applied Information Technology.
5. Effective Data Mining Using Neural Networks, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 8, NO. 6, DECEMBER 1996 by Hongjun Lu, Member, IEEE Computer Society, Rudy Setiono, and Huan Liu.
6. Evolving Artificial Neural Networks, by XIN YAO, PROCEEDINGS OF THE IEEE, VOL. 87, NO. 9, SEPTEMBER 1999.
7. Principal Feature Classification by Qi Li and Donald W. Tufts, IEEE TRANSACTIONS ON NEURAL NETWORKS, VOL. 8, NO. 1, JANUARY 1997.
8. Neighborhood size selection in the k-nearest neighbor rule using statistical confidence, JigangWang, Predrag Neskovic, Leon N. Cooper.
9. Tom Mitchell, Machine Learning, McGraw-Hill Computer science series, 1997.