# Performance Enhancement in Machine Learning System using Hybrid Bee Colony based Neural Network

**S. Karthick [1*]**

[1] *Team Manager, Sea Sense Softwares (P) Ltd., Marthandam, Tamil Nadu, India*

---

**Abstract -** *Metric learning and data prediction is a notable issue in many data mining and machine learning applications. In recent advances many researches have been proposed for the data prediction. Still there is a problem such as reduced prediction accuracy, computational time cost etc. Hence in this paper a novel Hybrid Bee Colony based Neural Network (HBCNN) is proposed for the data prediction in data mining. The proposed technique is segmented into three phases, in the first phase the raw input data is pre-processed to make it suitable for computation. The second phase is training of the proposed HBCNN technique is which the neural network is trained using bee colony algorithm. The last phase is dynamic testing in which the neural network is undergone rapid testing with dynamic input for the validation and testing of trained ANN. The proposed technique is tested using cancer dataset and the performance is compared with KNN technique.*

***Key Words*:** **Metric Learning, neural network, bee colony, data prediction**

## 1.INTRODUCTION

Artificial Neural Network (ANN) is often a new industry connected with computational science which integrates the various strategies for difficulty resolution which cannot be consequently effortlessly described without the help of an algorithmic conventional concentration. The particular ANNs stand for big as well as diverse classes connected with computational types [1]. Together with biological template modules these people are created basically by more or less comprehensive examples. The particular general approximates as well as computational types having specific traits such as the ability to learn or even adapt, to prepare or to generalize data are recognized as the particular Artificial Neural Networks. Contemporary development of a (near) optimal network architecture is carried out by means of people skilled together with use of a monotonous learning from trial and error process. Neural networks are generally algorithms regarding optimization as well as finding out, and are dependent freely with principles prompted merely by researching on the characteristics from the brain. For optimization as well as finding out neural networks in tandem with genetic algorithms, there are generally two techniques, each using its own advantages coupled with weaknesses. By means of different paths, the two processes are commonly evolved [2, 3].

Optimization algorithms, in addition to being called Learning algorithms, are, according to a number of features of scientific advancement, usually known as Genetic algorithms. An easy method involving encoding answers to the issue in chromosomes, an assessment purpose which on return starts to attain Chromosome directed at the idea, involving initializing population of chromosomes, workers that could be put on parents whenever they reproduce to correct the genetic composition are the five expected criteria. Integration could be of mutation, crossover in addition to site-specific workers. Parameter configurations for the particular criterion are the workers or can be anything else [4]. The criteria can develop populations involving much better ones in addition to the much better individuals, converging eventually in effect close to an international ideal, every time a genetic algorithm is actually run having a portrayal which usefully encodes a solution to a problem in addition to workers that can generate much better young ones through good parents [5]. In numerous situations, with regard to performing the particular optimization on the standard workers, mutation in addition to crossover, is usually adequate [6]. Genetic algorithms can assist like a black box function optimizer certainly not requesting any kind of know-how about computers of the particular domain in these cases. However, understanding of the particular domain is frequently exploited to raise the particular genetic algorithms performance with the incorporation involving new workers highlighted in this particular paper [7].

In neural network, functionality enhancement partition space can be a space which is used to classify data sample right after the test is actually mapped by neural network [8]. Centroid can be a space with partition space and also denotes the particular middle of the class. In traditional neural network functionality, location of centroids and the partnership involving centroids and also classes are generally fixed personally [9]. Furthermore, with reference to the quantity of classes variety of centroids is actually set. To locate optimum neural network this particular set centroid restriction minimizes the risk [9]. Genetic algorithms (GAs) have emerged to be practical software for the heuristic alternative of sophisticated discrete optimization issues. Particularly, there has been large fascination with their use in the most effective arranging and timetabling issues [10]. However, nowadays there have been several attempts made to become listed on both systems. Neural networks could be rewritten since a type of genetic algorithm is termed as some sort of classifier program and also vice-versa [11]. Genetic algorithm is meant for training feed forward networks. It does not merely work with its task but it does perform rear distribution, the normal training criteria [12, 13]. This accomplishment comes from tailoring the particular genetic algorithm to the domain of training neural networks.

Although the ANN features enjoy several advantages it has got several complications similar to convergence in addition to receiving throughout neighborhood minima during training by means of back propagation [14]. That is why the analysts are engaged to produce a new finest protocol to train the ANN. On this routine optimization protocol, algorithms similar to PSO, CS etc have been utilized to enhance the overall performance however it certainly does not fulfill the overall performance prerequisites. On this impression our system is also thought out to be able to get over a real issue during training involving ANN. Thus, we proudly propose the novel AGCS technique for training the ANN.

## 2. PROPOSED HBCNN TECHNIQUE FOR MACHINE LEARNING SYSTEM

Metric learning and Data prediction is a notable issue in many data mining and machine learning applications. In recent advances many researches have been proposed for the data prediction. Still there is a problem such as reduced prediction accuracy, computational time cost etc. Hence a novel hybrid HBCNN is proposed for the data prediction in data mining. The architecture of the proposed system is given in fig 1.

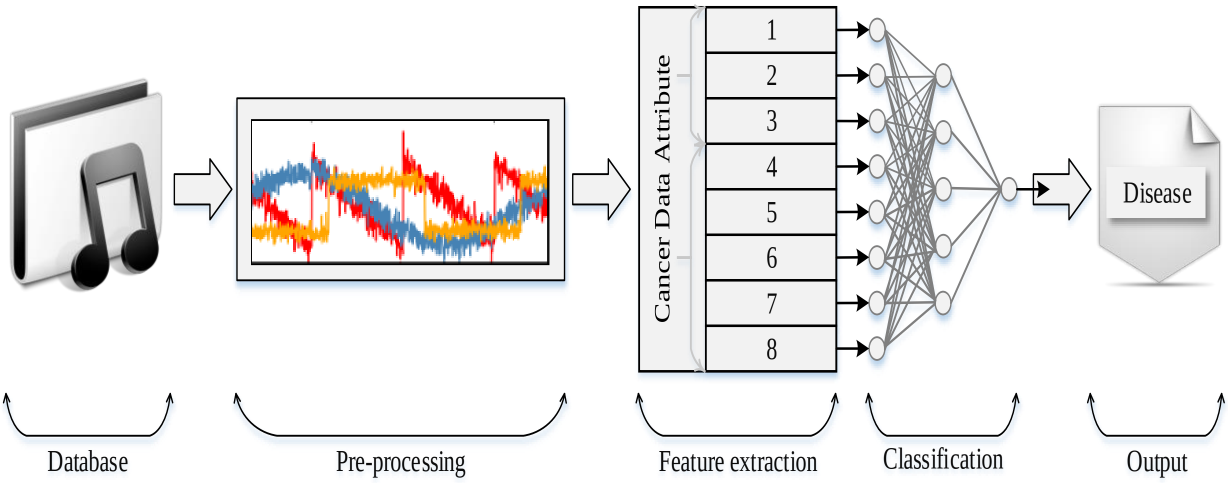


**Fig -1:** Architecture of the proposed system

The proposed system includes four phases they are;

1. Data Collection
2. Data Preprocessing
3. ANN-ABC Predictor Modelling
4. Dynamic Testing

### 2.1 Artificial Neural Network

An ANN is a programmed computational model that aims to duplicate the neural structure and the functioning of the human brain. It is made up of an interconnected structure of artificially produced neurons that function as pathways for data transfer. Artificial neural networks are flexible and adaptive, learning and adjusting with each different internal or external stimulus. Artificial neural networks are used in sequence and pattern recognition systems, data processing, robotics and modeling. The ANN consists of a single input layer, and a single output layer in addition to one or more hidden layers. All nodes are composed of neurons except the input layer. The number of nodes in each layer varies depending on the problem. The complexity of the architecture of the network is dependent upon the number of hidden layers and nodes. Training an ANN is to find a set of weights that would give desired values at the output when presented with different patterns at its input. The two main process of an ANN is training and testing.

The total of eight attributes or input is given as the input for the ANN and the corresponding class or disease can be obtained as the output of the ANN. Thus the proposed ANN architecture contains eight inputs (eight attributes) and corresponding disease output. The main two process of a classification algorithm is training and testing.

In training the input as well as the output will be defined and the appropriate weight is fixed so that the classifier (ANN) able to predict the apt object (disease) in the testing phase. Hence the training phase is the major part in a common classifier algorithm. In the proposed system the artificial bee colony (ABC) algorithm is used for the training. The process involved in the training is given as follows and the architecture of the back propagation neural network (BPNN) is given in the fig 2.

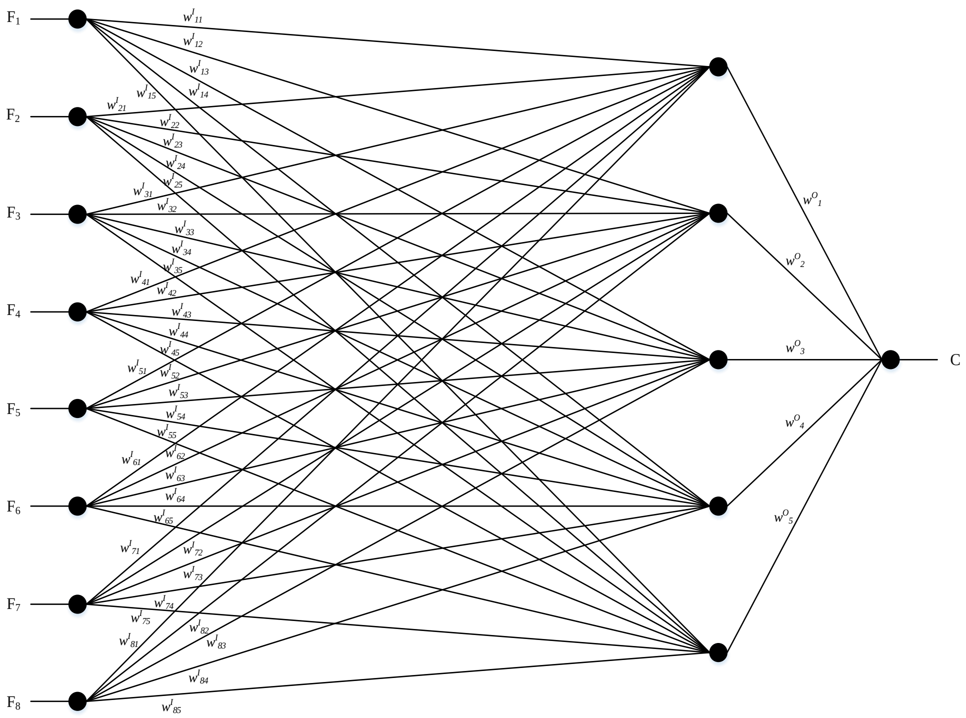


**Fig -2:** Proposed back propagation neural network

The proposed ANN consist of eight input units, one output units, and $M$ hidden units (M=5). First, the input data is transmitted to the hidden layer and then, to the output layer. This is called the forward pass of the back propagation algorithm. Each node in the hidden layer gets input from the input layer, which are multiplexed with appropriate weights and summed. The output of the neural network is obtained by the Eqn. (1) given below.

$$C = \sum_{j=1}^{M} \frac{w_j^O}{1 + \exp(-\sum_{i=1}^{N} F_i w_{ij}^I)} \qquad (1)$$

In Eqn. (1), '$F_i$' is the $i^{th}$ input value and '$w_j^O$' is the weights assigned between hidden and output layer, '$w_{ij}^I$' is the weight assigned between input and hidden layer and M is the number of hidden neurons. The output of the hidden node is the non-linear transformation of the resulting sum. Same process is followed in the output layer. The output values from the output layer are compared with target values and the learning error rate for the neural network is calculated, which is given in Eqn. (2).

$$fit = \partial_k = \frac{1}{2}(Y - C)^2 \qquad (2)$$

In Eqn. (2), $\partial_k$ is the $k^{th}$ learning error of the ANN, Y is the desired output and C is the actual output. The error between the nodes is transmitted back towards the hidden layer. This is called the backward pass of the back propagation algorithm. Then the training is repeated for some other training dataset by changing the weights of the neural network. The learning error is considered as the fitness function in ABC for the error minimization.

### B. Artificial Bee Colony Algorithm

Artificial Bee Colony (ABC) algorithm is a comparatively new technique proposed by Karaboga. ABC is encouraged by the foraging behaviour of honey bee swarms. In the ABC algorithm, the colony consists of three kinds of bees namely:

- Employee bee
- Onlooker bee
- Scout bee

Among these three kinds of bees, the employee bee and the onlooker bee are the employed bees, whereas, the scout bee is an unemployed bee. The number of employee bees and onlooker bees are said to be equal in ABC algorithm. The food sources are considered as the possible solutions for a given problem and the nectar amount of the food source is the relative fitness of that particular solution. The population of the colony is twice the size of the food sources. The number of food sources represents the position of the possible solutions of the optimization problem and the nectar amount of a food source represents the quality (fitness) of the associated solution. The pseudo code of the ABC algorithm is given below.

**Initialize food sources**

**Repeat**

**Employed Bee Phase:**

Send the employed bees to the food sources assigned to them.

Determine the amount of nectar (fitness value) in the food source.

Calculate the probability value of the food sources.

**Onlooker Bee Phase:**

Select the food sources discovered by the employee bee based on the probability value.

Find the neighboring food source and estimate its nectar amount.

Compare both the food sources and select the food source with better fitness.

**Scout Bee Phase:**

Select the food sources randomly and replace the abandoned food source with the new food source.

Memorize the best food source.

**Until (Requirements are met)**

In the initial step of ABC algorithm, a set of food sources are selected randomly by the employed bee and their corresponding nectar amounts (i.e. fitness value) are computed. The employee bee shares these details to another set of employed bees called as onlooker bees. After sharing this information, the employee bee visits the same food source again and then, finds a new food source nearby using the Eqn. (3).

$$v_{i,j} = x_{i,j} + \Phi_{i,j}(x_{i,j} - x_{k,j}) \quad (3)$$

Where, k and j are random selected index that represent the particular solution from the population. $\Phi$ is a random number between [-1,1].

When new neighbouring food sources are generated, their fitness values are calculated and the employee bee applies greedy selection to make a decision on whether to replace the existing food source in memory using new food source or not. Then, the probability value $P_i$ is calculated for each food source based on its fitness amount using the following Eqn. (4).

$$P_i = \frac{fit_i}{\sum_{i=1}^{CS/2} fit_i} \quad (4)$$

*Where, CS* is the colony size.

## 3. Results and Discussion

The proposed system is tested using cancer dataset and the results are compared with conventional techniques. The ALL / AML dataset for this experimental analysis is collected from online. After the normalization the randomly chosen sample is divided into three categories such as training, cross validation and testing data sets. The training data set is used for learning the network. Cross validation is used to measure the training performance during the training as well as to stop the training if necessary. The Leukaemia cancer contains four types. In our system we consider only two major types namely ALL (Acute Lymphoblastic Leukaemia) and AML (Acute Myeloid Leukaemia). For this training purpose we consider 38 patients who are divided into two clusters with 25 and 13 patients for ALL and AML respectively with overall 7129 genes.

The performance of the proposed system is compared base on the execution time and accuracy of the prediction. Execution time is the time in which a single instruction is executed. It makes up the last half of the instruction cycle. Accuracy is used to describe the closeness of a measurement to the true value. When the term is applied to sets of measurements of the same measured, it involves a component of random error and a component of systematic error. In this case trueness is the closeness of the mean of a set of measurement results to the actual (true) value and precision is the closeness of agreement among a set of results.

**Table 1:** performance analysis table for k=1

| K=1 | Techniques | | | | | |
|---|---|---|---|---|---|---|
| | Proposed | | Existing System | | Cosine Based System | |
| Data Set | accuracy | execution time | accuracy | execution time | accuracy | execution time |
| iris | 96.00% | 0.148 seconds. | 96.00% | 0.16 seconds. | 94.67% | 0.022 seconds |
| glass | 71.03% | 0.069 seconds. | 67.29% | 0.11 seconds. | 67.29% | 0.023 seconds. |
| vowel | 96.77% | 0.174 seconds. | 96.36% | 0.51 seconds. | 91.72% | 0.119 seconds. |
| cancer | 67.14% | 0.124 seconds. | 63.59% | 0.88 seconds. | 65.48% | 0.116 seconds. |
| letter | 93.95% | 3.569 seconds. | 94.55% | 4.878 seconds. | 94.08% | 9.692 seconds. |
| DNA | 73.38% | 1.886 seconds. | 73.38% | 1.938 seconds. | 71.25% | 3.013 seconds. |

For k=1 performance analysis is compared with the existing technique. For the cosine based system execution time is evaluated with the proposed and the existing strategies. Comparison analysis is evaluated for the existing and the cosine based system. This comparison is carried out for the iris, glass, vowel, cancer, letter and DNA.

For k=2 performance analysis is compared with the existing technique. For the cosine based system execution time is evaluated with the proposed and the existing strategies.

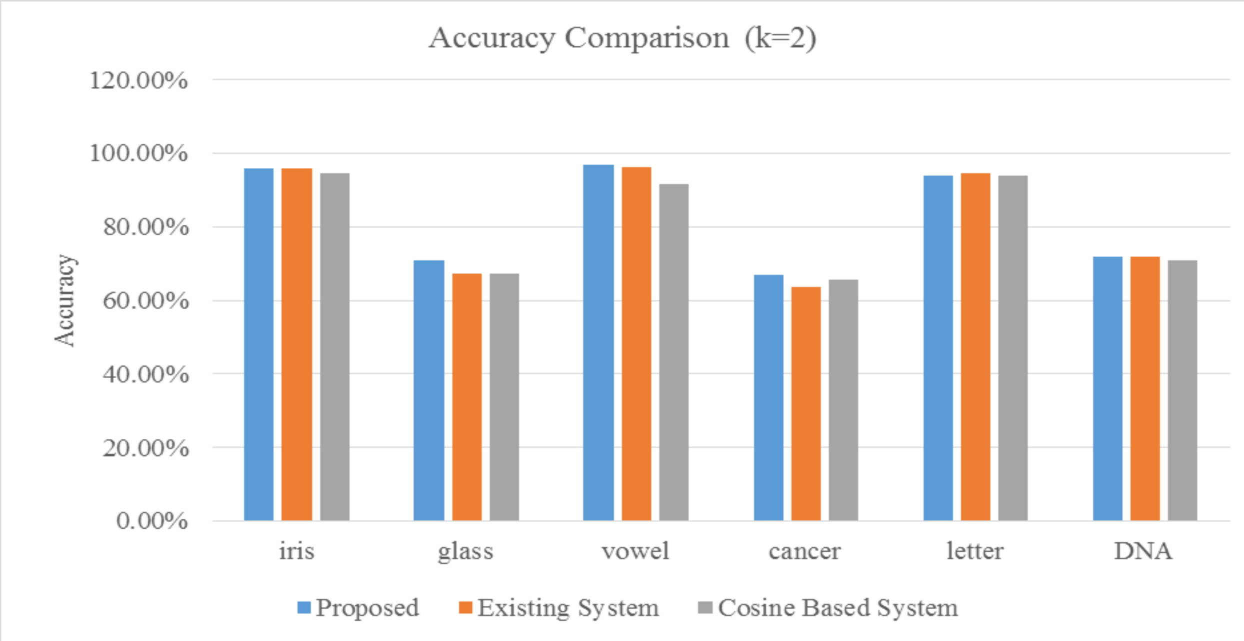


**Fig -3:** Performance analysis chart for k=2

Figure 3 shows the performance analysis of the chart for k=2.Comparison analysis is evaluated for the existing and the cosine based system. This comparison is carried out for the iris, glass, vowel, cancer, letter and DNA.

Figure 4 shows the performance analysis of the chart for k=1 for the execution time. Comparison analysis is evaluated for the existing and the cosine based system. This comparison is carried out for the iris, glass, vowel, cancer, letter and DNA.

For k=3 performance analysis is compared with the existing technique. For the cosine based system execution time is evaluated with the proposed and the existing strategies.

Figure 5 shows the performance analysis of the chart for k=3. Comparison analysis is evaluated for the existing and the cosine based system. This comparison is carried out for the iris, glass, vowel, cancer, letter and DNA.

Figure 6 shows the performance analysis of the chart for k=1 for the execution time. Comparison analysis is evaluated for the existing and the cosine based system. This comparison is carried out for the iris, glass, vowel, cancer, letter and DNA.

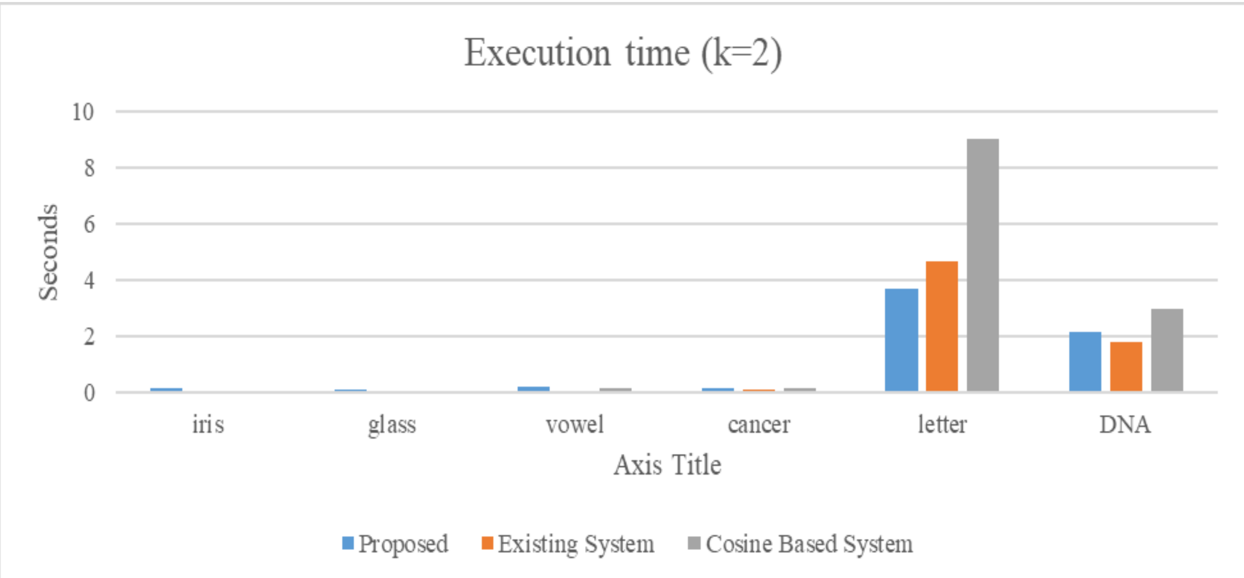


**Fig -4:** Performance analysis chart the execution time for K=2

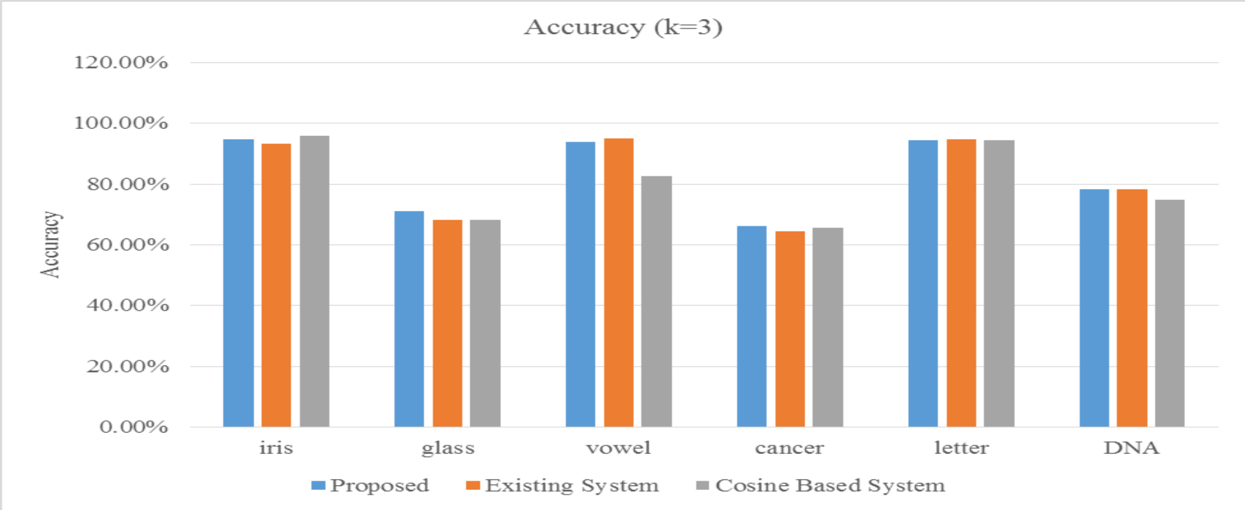


**Fig -5:** Performance analysis chart for k=3

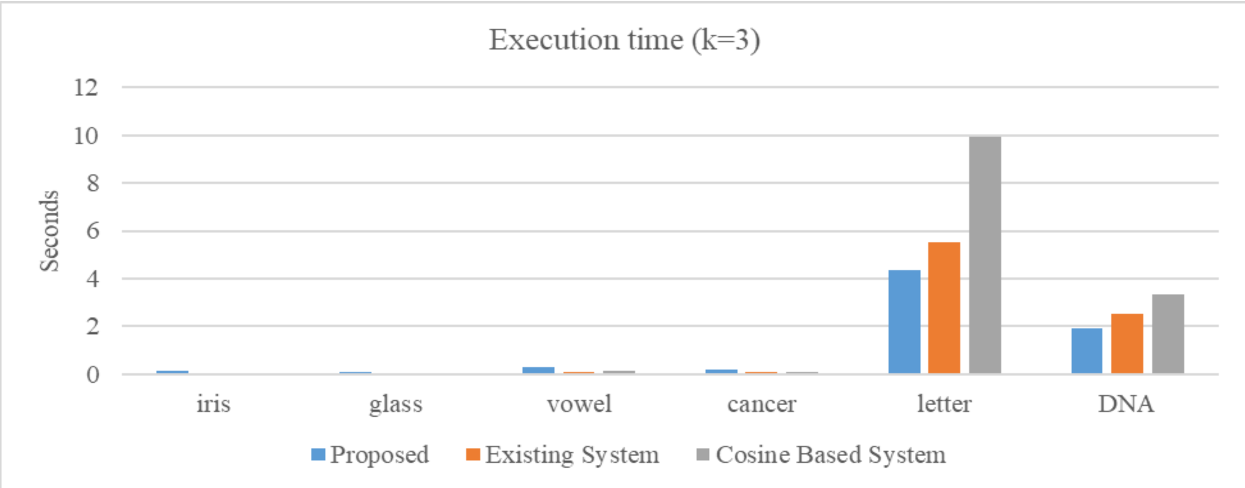


**Fig -6:** Performance analysis chart the execution time for K=2

## 4. Conclusion

Machine learning is the motivated research topic in the recent decade. Metric learning become a challenging task in machine learning. Hence a novel procedure for the metric learning is proposed in this paper. The proposed system uses two stages of execution, in the first stage semi supervised clustering and second stage prediction is performed. The hierarchy forest clustering technique is utilized for the semi supervised clustering and KNN technique for the prediction. The performance of the system is verified based on the accuracy and execution time. The performance analysis outperforms the existing techniques and proves its effectiveness for the metric learning in machine learning.

## REFERENCES

[1] Mantas Lukosevicius, and Herbert Jaeger, "Reservoir Computing Approaches to Recurrent Neural Network Training," Computer Science Review, Vol. 3, 2009, pp. 127–149.

[2] David E. Moriarty, Alan C. Schultz, and John J. Grefenstette, "Evolutionary Algorithms for Reinforcement Learning," Journal of Arterial Intelligence Research, Vol. 11, 1999, pp. 241-276.

[3] Lucas Negri, AdemirNied, Hypolito Kalinowski, and Alexander Paterno, "Benchmark for Peak Detection Algorithms in Fibre Bragg Grating Interrogation and a New Neural Network for its Performance Improvement," Sensors, Vol. 11, 2011, pp. 3466-3482,

[4] David J. Montana and Lawrence Davis, "Training Feed forward Neural Networks Using Genetic Algorithms," In Proceedings of the 11th international joint conference on Artificial intelligence, Vol. 1, 1989, pp. 762-767,

[5] William B. Langdon, Robert I. McKay and Lee Spector, "Genetic Programming," International Series in

Operations Research & Management Science, Vol. 146, 2010, pp. 185-225,

[6] Swagatam Das, Ajith Abraham, Uday K. Chakraborty, and Amit Konar, "Differential Evolution Using a Neighborhood-Based Mutation Operator," IEEE Transactions on Evolutionary Computation, Vol. 13, 2009, pp. 526-553,

[7] Adam M. Smith and Michael Mateas, "Answer Set Programming for Procedural Content Generation: A Design Space Approach," IEEE Transactions on Computational Intelligence and AI in Games, Vol. 3, 2011, pp. 187 -200,

[8] Lin Wang, Bo Yang, Yuehui Chen, Ajith Abraham, Hongwei Sun, Zhenxiang Chen and Haiyang Wang, "Improvement of neural network classifier using floating centroids," Vol. 31, 2012, pp. 433-454,

[9] [9] Lei Zhang, Lin Wang, Xujiewen Wang, Keke Liu, and Ajith Abraham, "Research of Neural Network Classifier Based on FCM and PSO for Breast Cancer Classification," Hybrid Artificial Intelligent Systems, Vol. 7208, 2012, pp. 647-654,

[10] Tanomaru J., "Staff Scheduling by a Genetic Algorithm with Heuristic Operators," In Proceedings of the IEEE Conference on Evolutionary Computation, 1995, pp. 456-461.

[11] Patrycja Vasilyev Missiuro, "Predicting Genetic Interactions in Caenorhabditis Elegans using Machine Learning," Massachusetts Institute of Technology, 2010, pp. 191-204

[12] Piotr Mirowski, Deepak Madhavan, Yann LeCun, and Ruben Kuzniecky, "Classification of patterns of EEG synchronization for seizure prediction," Clinical Neurophysiology, Vol. 120, 2009, pp. 1927–1940

[13] XS Yang and S. Deb, "Cuckoo search via Lévy flights," In Proceedings of World Congress on Nature & Biologically Inspired Computing, 2009, pp. 210-214

[14] L. Mingguang and L. Gaoyang, "Artificial Neural Network Co-optimization Algorithm based on Differential Evolution," In Proceedings of Second International Symposium on Computational Intelligence and Design, 2009