

Population Based Meta Heuristics Algorithm for Performance Improvement of Feed Forward Neural Network

Dr. Samuel Manoharan,
Professor,
Department of Electronics,
Bharathiyar College of Engineering and Technology,
Thiruvettakudy, Karaikal, India
Email id: jsamuel@bcetedu.in

Prof. Sathish
Department of EEE,
Eritrea Institute of Technology,
Eritrea
Email id: sathesh4you@gmail.com

Abstract: The most vital step in mining data's in order to have a proper decision making is the classification, it is remains important in multiple of human activities such as the industrial applications, marketing campaigns, research process and the scientific endeavors. The process of classifying involves the objects categorization into classes that are already defined. These categorizations are developed according to the identical attributes of the items or the objects. Multitudes of methods were devised to improve the accuracy in the classification to devour an enhanced performance in terms of faster convergence speed. The algorithm based on water cycle that includes the evaporation, condensation and precipitation (WC-ECP), which is a population based metaheuristic is used in the paper to improve the accuracy in the feed forward neural network (PNN-probabilistic neural network) to standardizes its random constraint choice and in turn improvise the accuracy of the categorization and the speed of the convergence. The proposed method was tested with the five dataset of UCI machine learning repository and was evinced that the WCECP-PNN performed better compared to the other evolutionary algorithms such as the GA which is also a population based Meta-heuristics

Keywords: Data Mining, Classification, Feed Forward Networks, PNN, WCECP

1. Introduction

Data mining is the science of extracting useful knowledge from vast repositories in many fields, such as industry, academic research and medical practice, using automated search processes that employ statistical and computational techniques. It includes the detection of relevant patterns and the automated analysis and exploration of broad data sets by quantities to recognize secret relationships this is used in forecasts where

certain contents are used to estimate remaining contents (categorize) or in explanations where patterns can be easily interpreted by the user (clustering) are defined.

Categorization or distinguishing objects based on its similar attributes is a type of supervised-ML which is widely used in various fields such as science, medicine, business, and industry to support decision-making. Challenges in classification occur when an object must be categorized based on a variety of object attributes. The most critical step in the categorization process is to choose the correct procedure which can apply to all real-world issues. The classification methods are available to use only after primary testing has been completed and the tests have been considered appropriate.

Researchers have developed many techniques focused on artificial intelligence which may include artificial neural networks (ANNs), KNN , SVM, naive Bayes classifier, RBF iterative dichotomiser 3 (ID3) algorithm and many more. The neural networks utilizes the heuristic models to solve the class statistics, which repeatedly modifies the system parameters to boost the system efficiency. Subsequently, during preparation, this technique expends a lot of computation time, which can lead to false minima. The PNN was developed to disable or reduce this shortcoming of the methodology. The PNN is feed-forward-NN which can be used to address nonlinear issues. It is particularly possible because in PNN, the NN's sigmoid activation function is replaced by an exponential function derived from a Bayesian decision rule. The PNN structure consists of four layers: the layer of data, the layer of sequence, the summation layer and finally the layer of output it has the ability to achieve maximum ranking, and is more reliable than NNs. However, a PNN has its limitations— for example, detection of new objects or instances it is deliberate than an NN and necessitates supplementary resources for computation, e.g. Memory space. Scientists have hybridized different categorization methods, where the positives or advantages of the individual algorithms have been combined in order to achieve better results. Therefore, with the aid of metaheuristic algorithms, the shortcomings of the classification techniques have been reduced to some degree, and thus the performance of the categorization techniques are improved.

The most important goal behind the incorporation of two algorithms is to merge both the algorithm strengths to evade their specific restrictions. Each metaheuristic algorithm listed here has its own domain of success. The GA algorithm, for example, is very successful when it comes to exploration but is not useful for exploitation. So in the proposed method examines the performance enhancement achieved in the PNN in distinguishing by utilizing the WCECP algorithm that is based on the water evaporation cycle. In the proposed process the PNN develops a random solution initially and attains enhancement in the performance by employing the WCECP to improvise the PNN weights by successfully exploring, exploiting and avoiding convergence that is premature as well as population immobility which in turn increases the capability of the PNN categorization to identify a solution that is optimal.

The paper integrating the PNN with the WCECP to attain a solution that optimal for the issues in categorization is elaborated with the related works is 2, the basic description of PNN-FFNN and WC-ECP in 3. The proposed method that integrates the PNN with the WCECP. Results evaluation in 5 and Conclusion in 6.

2. Literature Survey

Specht et al [1] elaborates the tutorial on the "Probabilistic neural networks." Along with its classification capabilities. Mahajan et al [2] presents the "Review of data mining techniques for churn prediction in telecom." Andreopoulou, et al [3] puts forth the "methodologies of the clustering" Johnston proposes the "Generation of a supervised classification algorithm for time-series variable stars with an application to the LINEAR dataset." Hernández-Lobato et al [5] proposes the "Probabilistic backpropagation for scalable learning of bayesian neural networks."

Schaffer et al [6] details the survey on integrating the genetic algorithm and neural networks. Smys, S. et al [7] discuss the utilization of machine learning in the detection of DDOS Sebt, M. H et al [8] presents the "Hybridization of genetic algorithm and fully informed particle swarm for solving the multi-mode resource-constrained project scheduling problem." Kumar et al [9] proposes the Pattern recall analysis of the Hopfield neural network with a genetic algorithm." Raj, Jennifer S et al [10]puts forth the "A Comprehensive Survey on the Computational Intelligence Techniques and Its Applications." Cruz et al [11] presents the. "A bee-inspired data clustering approach to design RBF neural network classifiers." Adio et al [12] proposes the "Experimental investigation and model development for effective viscosity of MgO–ethylene glycol nanofluids by using dimensional analysis, FCM-ANFIS and GA-PNN techniques." Manimala, K et al [13] puts forth "Power disturbances classification using S-transform based GA–PNN." Joseph, S et al [14] elaborates the "Survey of data mining algorithms for intelligent computing system."

3. Back Ground

3.1. Probabilistic-NN Architecture

The architecture of P-NN is consists of three planes (layers), with a layer for the input in the left comprising 'L' number of nodes to represent every 'L' input feature of the feature vector. The output from the input layer is fed as the input to the hidden layer that is in the middle such that every node in the hidden layer take delivery of the entire input feature vector 'i'. Every node in the hidden layer is grouped as single cluster for every classes of 'C' types. The nodes of the hidden plane in the cluster for a class 'C' resembles the Gaussian function involved on its related feature vector of the class c, A Gaussian exists for every feature

vector that is exemplar. Every Gaussians in the Class of cluster input their functional values to the corresponding node at the output plane for that class, leaving ‘C’ nodes at the output. The figure.1 below shows the P-NN architecture with the capability to recognize the C = 2 classes, meanwhile this architecture could also be extended to any number of classes.

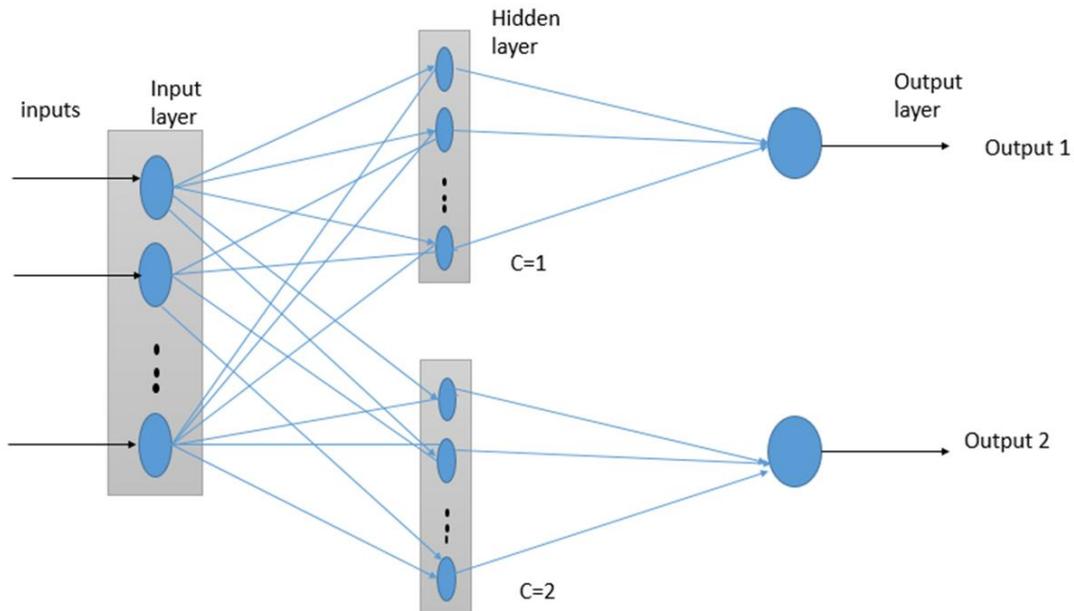


Figure.1 PNN Architecture

In the output of the nodes in the third layer that is the output plane , adds all the Gaussian values observed for a particular class of classes ‘c’ and scales the added results such that the probability of the sum function is unity , forming a probability density function. Unique identities are utilized temporarily to improve the clarity. The exemplar feature vectors for C=1 is assigned as ‘X’ and C =2 is assigned as ‘Y’ where $(N^X : x = 1 \dots X)$ and $(M^Y : y = 1 \dots Y)$, so the layer that is hidden holds X and Y nodes for the C=1 and C=2 respectively, the equations 1 and 2 are Gaussian equations for the C =1 and 2 for any ‘l’ input vectors.

$$G_1(N) = \left(\frac{1}{\sqrt{2\pi\sigma^2}l} \right) e^{-|N-N^X| \wedge 2 / 2\sigma^2} \quad (1)$$

$$G_2(M) = \left(\frac{1}{\sqrt{2\pi\sigma^2}l} \right) e^{-|M-M^Y| \wedge 2 / 2\sigma^2} \quad (2)$$

Where the σ represents the one half of the average distance across the features vectors that are present in the same clusters or it could also be the distance of one exemplar to the next that is at proximity. The output node of the class 'C' sums all the corresponding values of cluster 'c' from the hidden node, this termed as the Parzen windows or the mixed Gaussians. The process does not have any iterations or computations for the estimation of weights, in case of the huge set of sum of Gaussians, the occurrence of error might be significant. For this the σ is made larger by reducing the size the feature vectors placed closely. The high level algorithm for the PNN given below makes the training set using the exemplar features, and assumed to have the knowledge of the class the each vector belongs to.

- a) Read exemplar vector file and class number.
- b) Sort into 'C' set such that each set contains a single vector class.
- c) \forall 'C' describe Gaussian function for the exemplar vector and the added values of Gaussian output function.
- d) Define PNN and utilize the Feed vector to categorize.
- e) Read input vector, feed to every Gaussian function (GF) belonging to each class.
- f) \forall hidden node clusters, estimate the GF
- g) \forall Hidden node clusters feed GF values to the one node that is set as output.
- h) \forall node at output representing each class add all inputs and multiply the constants.
- i) Identify the added function maximum values in the nodes set at the output.

3.2. WC-ECP Algorithm

The Algorithm is a population based metaheuristic that is based on the population of the water droplets that evaporates from the earth, to be condensed, precipitated and formed as rain this process was introduced by the Sadollah et al in the year 2012 The following steps describe the implementation of the WC-ECP algorithm.

- i) Initialize the input parameters, the dimension of the population (p_D), the number of iterations (Max_{it}) and the parameter that is defined by the user (S_{max})
- ii) Initialize the population of the water droplet sources from the earth, (rivers (r), streams (st), ponds(p) and sea (s))
- iii) Estimate the cost for the water droplets Cost (r^i, st^i, p^i, s^i) where $i = 1, 2, 3 \dots \dots p_D$.

- iv) Define the intensity of the water source opted

- v) Exchange the positions of the optimal water sources with previous available solution.

- vi). Identify the evaporation condition, by taking mod for the value obtained by subtracting current optimal water source and the existed source and determine the solution such that it is less than S_{max}

- vii) Detect evaporating conditions denoted by $optimal\ source(t + 1) = LB + (rand * (UB - LB))$, where the UB and the LB are the upper and the lower bounds respectively.

- viii). Minimize the value of S_{max} such that the $S_{max}(t + 1) = S_{max}(t) - \left[\frac{S_{max}(t)}{Max_{it}} \right]$

- ix). The convergence norms are measured if satisfied the process is terminated else goes to iii.

The WE-ECP enables to estimate the weights that are optimal and could be utilized in the Probabilistic-NN algorithm. The procedure initiates by randomly generating the initial weights for the PNN. The values at the input are multiplied with the associated weights (W_{xy}) according to the values evaluated by the model of the P-NN.

4. Proposed W-ECP and PNN Architecture

The proposed structure is encompassed of two sections one the probabilistic-neural network that does the training and the testing to perform the categorization followed by the next section that is the WC-ECP where the weights are adjusted in order to improve the accuracy. The repeated procedure is stopped when the stopping norms are met. The accuracy of the P-NN by clubbing the W-ECP and PNN is enhanced by considering both the exploration as well as exploitation. In the exploitation phase of the W-ECP the optimal sources of the water droplets are estimated and the current sources and their locations are replaced with the optimal sources and their corresponding locations respectively. The early convergence is evaded by the constant checks performed on the evaporating conditions and the estimation of the optimal sources. The exploration phase depends on the populace initialization, it utilizes a heuristic approach and produces the population in greedy manner and based on the current minimum, estimating the cost value for the source

of water available, to preventing the stagnation of the population and elude the early stage convergence in turn causing a balance between the exploration and the exploitation of the search space. The flow chart below shows the fine tuning process of the PNN weights to achieve enhanced accuracy in the categorization.

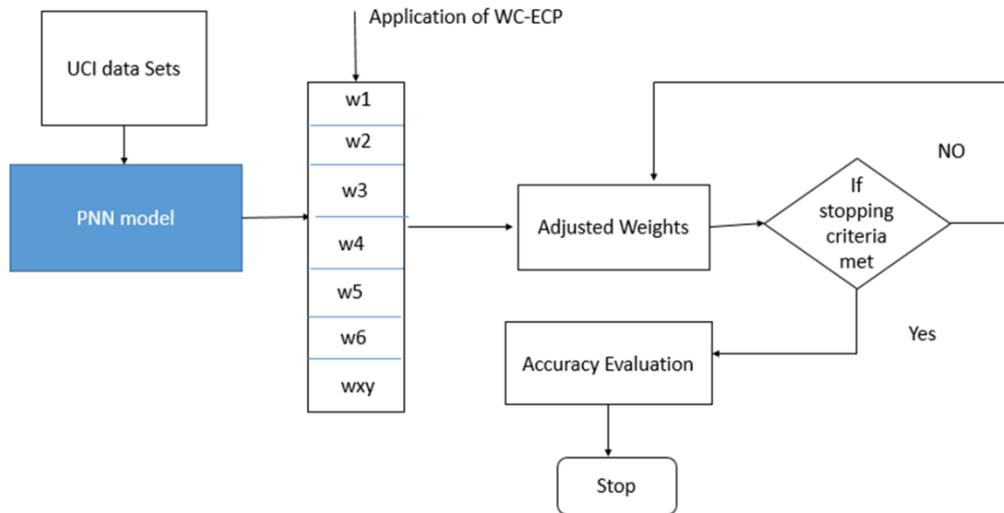


Figure.2 PNN Weight Adjustment

5. Result Evaluation

The PNN Structure Hybridized with the WC-ECP algorithm to adjust the weights of the PNN is validated using the 5 dataset of UCI machine learning source that could be downloaded freely. For every data set 5 runs were performed independently according to the procedures provided in the website. The list of the data used and the parameter setting are described in the table. 1 and 2 below

Dataset	Attributes Count	Training Data set	Testing Data set
Breast Cancer	10.01	194	73
Parkinson's	24	132	50
Disorders of liver	6.05	232	87
PIMA Indian Diabetes	8.001	516	194
Heart	13.01	184	66

Table .1 Data set Used

Parameter	Values
Number of class	2
Number of water sources	4
Number of data set	5
Total Distance covered	$1e^{-16.01}$
Initial population	100
UB and the LB	1 and 0

Table. 2 Parameter Used

The true positive (t_p), false positive (f_p), true negative (t_n), and false negative (f_n), are utilized to determine the accuracy. Further the sensitivity, accuracy, specificity, and rate of error was estimated for the PNN clubbed WCECP the table.3 below shows the accuracy in the categorization along with the sensitivity, specificity and the error rate observed.

Data set	Sensitivity	Specificity	Error rate	Accuracy
Breast Cancer	.8910	.5674	.245	85
Parkinson's	.765	.5784	.214	88
Disorders of liver	.9845	.734	.123	89
PIMA Indian Diabetes	.8905	.8902	.178	90
Heart	.7564	.9813	.1492	92

Table.3 Observed Accuracy, Specificity, Sensitivity and Error Rate

Where the accuracy is measured using the $[A/B]$, sensitivity is estimated applying $[t_p/C]$, and applying $[t_n/D]$, for specificity whereas $1 - [A/B]$ for evaluating the rate of error occurrence. Where $(t_p + t_n) = A$, $t_p + t_n + f_p + f_n = B$, $t_p + f_n = C$, and $t_n + f_p = D$

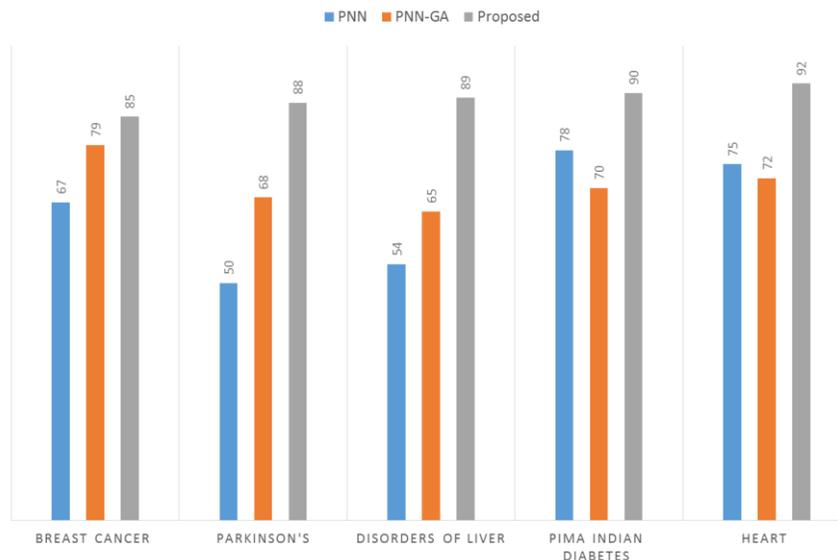


Figure. 3 Classification Accuracy

The figure.3 shows the classification accuracy observed for the different dataset , it was found the WC-ECP clubbed with the PNN provided better classification accuracy compared to the PNN with GA and the PNN alone the. The balance achieved in the exploration and the exploitation of the search space has made possible to enhance the accuracy in the PNN classification by adjusting the weights.

6. Conclusion

The methods to enhance the classification in the feed forward neural networks was put forth in the paper to have a quality suggestions that could be employed in the problem of categorizations. For this purpose the PNN was clubbed with the WC-ECP metaheuristic optimization. This WC-ECP with the balance achieved in the exploration and the exploitation of the search space fine-tuned the weights of the PNN and enhances the accuracy of the categorization. Moreover the PNN clubbed WC-ECP was tested on five datasets of UCI repository and the results were found convincing with the improvement attained in the accuracy of the classification.

References

- [1] Specht, Donald F. "Probabilistic neural networks." *Neural networks* 3, no. 1 (1990): 109-118.
- [2] Mahajan, Vishal, Richa Misra, and Renuka Mahajan. "Review of data mining techniques for churn prediction in telecom." *Journal of Information and Organizational Sciences* 39, no. 2 (2015): 183-197.
- [3] Andreopoulou, Zacharoula, Christiana Koliouka, and Constantin Zopounidis. "Methodologies." In *Multicriteria and Clustering*, pp. 13-32. Springer, Cham, 2017.
- [4] Johnston, Kyle B., and Hakeem M. Oluseyi. "Generation of a supervised classification algorithm for time-series variable stars with an application to the LINEAR dataset." *New Astronomy* 52 (2017): 35-47.
- [5] Hernández-Lobato, José Miguel, and Ryan Adams. "Probabilistic backpropagation for scalable learning of bayesian neural networks." In *International Conference on Machine Learning*, pp. 1861-1869. 2015.
- [6] Schaffer, J. David, Darrell Whitley, and Larry J. Eshelman. "Combinations of genetic algorithms and neural networks: A survey of the state of the art." In *[Proceedings] COGANN-92: International Workshop on Combinations of Genetic Algorithms and Neural Networks*, pp. 1-37. IEEE, 1992.
- [7] Smys, S. "Ddos Attack Detection In Telecommunication Network Using Machine Learning." *Journal of Ubiquitous Computing and Communication Technologies (UCCT)* 1, no. 01 (2019): 33-44.

- [8] Sebt, M. H., M. R. Afshar, and Y. J. E. O. Alipouri. "Hybridization of genetic algorithm and fully informed particle swarm for solving the multi-mode resource-constrained project scheduling problem." *Engineering Optimization* 49, no. 3 (2017): 513-530.
- [9] Kumar, Somesh, and Manu Pratap Singh. "Pattern recall analysis of the Hopfield neural network with a genetic algorithm." *Computers & Mathematics with Applications* 60, no. 4 (2010): 1049-1057.
- [10] Raj, Jennifer S. "A Comprehensive Survey On The Computational Intelligence Techniques And Its Applications." *Journal of ISMAC* 1, no. 03 (2019): 147-159.
- [11] Cruz, Dávila Patrícia Ferreira, Renato Dourado Maia, Leandro Augusto da Silva, and Leandro Nunes de Castro. "A bee-inspired data clustering approach to design RBF neural network classifiers." In *Distributed Computing and Artificial Intelligence, 11th International Conference*, pp. 545-552. Springer, Cham, 2014.
- [12] Adio, Saheed Adewale, Mehdi Mehrabi, Mohsen Sharifpur, and Josua Petrus Meyer. "Experimental investigation and model development for effective viscosity of MgO–ethylene glycol nanofluids by using dimensional analysis, FCM-ANFIS and GA-PNN techniques." *International Communications in Heat and Mass Transfer* 72 (2016): 71-83.
- [13] Manimala, K., and K. Selvi. "Power disturbances classification using S-transform based GA–PNN." *Journal of The Institution of Engineers (India): Series B* 96, no. 3 (2015): 283-295.
- [14] Joseph, S. Iwin Thanakumar, and Iwin Thanakumar. "Survey of data mining algorithm's for intelligent computing system." *Journal of trends in Computer Science and Smart technology (TCSST)* 1, no. 01 (2019): 14-24.