Enhancement of Health Data Accuracy Using Neural Network Algorithm

[1]Gautam Prakash, [2]Dr. Dinesh Kumar Sahu

Department of Computer Science and Engineering
Sarvepalli Radhakrishnan University, Bhopal

ABSTRACT

Now a days data mining technique used in the field of medical diagonise of critical desesis and clinical data. the prediction of mining technique is major issue. For the enhancement of mining technique used various approach such as fuzzy logic, feature optimization and machine learning based classification technique. in this paper proposed RBF model baed classification technique for the prediction of cilinical data. the prediction rate of data is good in compression of perivious methods. For the validation and vrfication of proposed model used MATLAB software and very reputed dataet such as blood cancer, stomach.

Keywords:- RBF, ANN, Fuzzy System, ID3, CRBF.

INTRODUCTION

ANN has absolutely settled itself as the most effectively displaying system for biomedical applications, especially in the range of example acknowledgment. Throughout the years, a few sorts of ANN have been produced, each with one of kind properties that make them more appropriate for certain errand over the others. The system engineering shifts as far as structure, actuation capacity and learning calculation. As a rule, the ANN can be executed in directed and unsupervised learning modes[4].

Here we discuss cascaded model of RBF network for ID3 classification. The great advantage of RBF network is single layer processing unit and target output independent with input data. During the time spent falling info highlight goes through edge of classifier, edge classifier work isolate information into layers, for example, positive and negative in information space area. The piece of positive and negative utilized as contribution as a part of fell model. We connected to prepare a neural system to take in the grouping highlights from the information tests of a minority class in the preparation set and to settle on more positive choices to the minority class. For the accommodation of portrayal, we alluded to the two classes of information as minority and larger part classes separately [10]. In numerous applications, if blunder is unavoidable, a neural system is relied upon to mistake on one specific class instead of the other the calculation we explored was to produce new minority information tests close to the arrangement limit utilizing the Gaussian and adds these new information tests to the preparation information. The neural systems prepared on this set ought to settle on more great choice to the minority class with the minimization of misclassification of the larger part class and have expanded speculation ability. For every data sample s of the minority class in the training set, we attempt to generate p new data samples around s subject to its local Gaussian distribution of the opposite class. Let us assume the input vector is M dimensional. In this section various methods discussed like support ID3, feature correlation and ID3 and finally discuss a proposed hybrid technique for ID-CRBF for classification.

Fuzzy rationale is a superset of routine (Boolean) rationale. It was first presented in the 1960s by Lofti A. Zadeh of the University of California, Berkeley as a way to handle the idea of incomplete truth (Zadeh 1965). Fluffy rationale is considered as a standout amongst the most effective instruments for managing imprecision and vulnerabilities. It was intended to scientifically speak to vulnerability and ambiguity and to give formalized apparatuses for managing the imprecision natural for some issues separated by the enrollment capacities (Coppin 2004)[8].
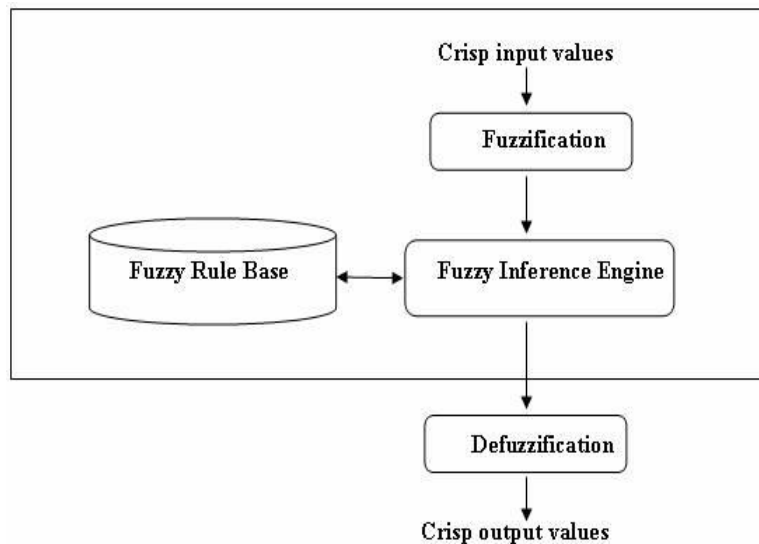
Figure 1: Fuzzy Logic Model.

A participation capacity is a bend that defines how every point in the info space is mapped to an enrollment esteem (or level of enrollment) somewhere around 0 and 1. The easiest participation capacities are shaped utilizing straight lines. Cases incorporate triangular participation work and trapezoidal enrollment work. On alternate hands, Gaussian and sigmoidal enrollment capacities are not made up of straight lines. Fluffy surmising is the way toward planning the mapping from an offered contribution to a yield utilizing fluffy rationale (Moraga 2000). The mapping defined gives the premise from which choices can be made, or designs perceived. The fluffy derivation process can be isolated into five sections which incorporate info fuzzification, applying fluffy administrators, applying suggestion strategies, yield total and finally yield defuzzification as delineated in Figure[8].

The rest of paper discuss as in section 2 discuss the Fuzzy Modeling. In section 3 discuss the ProblemFormulationt. In section 4 discuss proposed Work. In section 5 discuss the experimental result and analysis. finally discuss conclusion & future work in section 5.

2.FUZZY MODELING
The forecast undertaking talked about in this paper was performed utilizing fluffy displaying, because of the nonlinear way of much medicinal physiology, and many studies showing great appropriateness and execution with databases. An extra inspiration for utilizing fluffy demonstrating is the effectively justifiable tenets that are produced in the wake of arranging information, which is helpful for therapeutic understanding and rules creation[2].

Fluffy displaying is an instrument that permits an estimate of non-direct frameworks when there is practically no past learning of the framework to be demonstrated. A nitty gritty depiction of fluffy rationale and demonstrating can be found. Briefly, fluffy models utilize rules and coherent connectives to set up relations between the elements defined to infer the model. A fluffy classifier contains a lead base comprising of an arrangement of fluffy if-then standards together with a fluffy deduction system.
Since the relations between the info factors can have a nonlinear nature, fluffy frameworks were utilized as a part of this work to paired group pick up/no-pick up of data for every lab test as takes after: by utilizing variable Y as the lab test for which the data pick up is to be evaluated, and factors X as contributions to a fluffy model, X experience a forward choice of elements to acquire the subset of X that better classifies Y. To keep away from greatness impacts in the classification procedure, variable(s) X were standardized as follows[2]:

$$X_{norm} = \frac{(X - X_{min})}{(X_{max} - X_{min})}$$

where $Xnorm$ is the normalized version of X, while $Xmin$ and $Xmax$ speak to the base and most extreme estimations of X, individually. The minimum–maximum standardization strategy is ordinarily utilized as a part of designing applications to standardize the information because of its direct changing structure. Furthermore, Y was standardized by setting the lab tests with data pick up to 1, and those with no pick up to 0.

3. PROBLEM FORMULATION

For the purpose of medical disease data classification various machine learning algorithm are applied, such as clustering, classification, and regression. Two of the most critical and well generalized problems of medical data are its new evolved feature and concept-drift. Since a medical data is a fast and continuous event, it is assumed to have infinite length. Therefore, it is difficult to store and use all the historical data for training. The most discover alternative is an incremental learning technique. Several incremental learners have been proposed to address this problem [18], [7]. In addition, concept-drift occurs in the stream when the underlying concepts of the stream change over time. A variety of techniques have also been proposed in the literature for addressing concept-drift [12], [1], [7] in data stream classification. However, there are two other significant characteristics of data streams, such as concept evolution and feature evolution that are ignored by most of the existing techniques. Concept-evolution occurs when new classes evolve in the data. Cluster oriented ensemble classifier used to reduce feature evaluation problem in medical disease data classification. The selection of optimal number of cluster is also challenging job in medical disease data classification. On the review process we found some important problem in cluster oriented medical disease data classification. These problems are given below.

1.   Medical disease data classification suffered from new feature evaluation, concept evaluation ,feature evaluation and data drift[17]
2.   Selection of optimal cluster for ensemble classifier[1,3,6]
3.   Diversity of feature selection process.[12]
4.   Boundary value of cluster[9,13]
5.   Outlier data treat as noise [14, 15].

4. PROPOSED ALGORITHM

In this section we discuss cascaded model of RBF network for ID3 classification. The great advantage of RBF network is single layer processing unit and target output independent with input data. During the time spent falling info highlight goes through edge of classifier, edge classifier work isolate information into layers, for example, positive and negative in information space area. CRBF models are creating for data training for minority and majority class data sample for processing of tree classification. The input processing of training phase is data sampling technique for classifier. While single-layer RBF networks can potentially learn virtually any input output relationship, RBF networks with single layers might learn complex relationships more quickly. The function neID3 creates cascade-forward networks. For example, a cascaded layer network has connections from layer 1 to layer 2, layer 2 to layer 3, and layer 1 to layer 3. The cascade-layer network also has connections from the input to all cascaded layers. The additional connections might improve the speed at which the network learns the desired relationship. CRBF artificial intelligence model is similar to feed-forward back-propagation neural network in using the back-propagation algorithm for weights updating, but the main symptom of this network is that each layer of neurons related to all previous layer of neurons. Tan-sigmoid transfer function, log - sigmoid transfer function and pure linear threshold functions were used to reach the optimized status.

4.6.1 Process of Method

1.   Sampling of data of sampling technique
2.   Split data into two parts training and testing part
3.   Apply CRBF function for training a sample value
4.   Using 2/3 of the sample, fit a tree the split at each node

For each tree

➢   Predict classification of the available 1/3 using the tree, and calculate the misclassification rate = out of CRBF.

5.   For each variable in the tree
6.   Compute Error Rate: Calculate the overall percentage of misclassification

➢   Variable selection: Average increase in CRBF error over all trees and assuming a normal division of the increase among the trees, decide an associated value of feature.

7.   Resulting classifier set is classified

➢   Finally to estimate the entire model, misclassification

Decode the feature variable in result class

5. EXPERIMENTAL RESULT ANALYSIS

In this section we perform experimental process of proposed RBF method and ID3 algorithm. The proposed method implements in MATLAB 7.8.0 and tested with very reputed data set from UCI machine learning research center. In the research work, I have measured classification Sensitivity, Specificity,Accuracy and execution time of ensemble method. To evaluate these performance parameters I have used five datasets from UCI machine learning repository [10] namely blood dataset, diabetes dataset, harerman dataset, heart dataset and liver dataset.
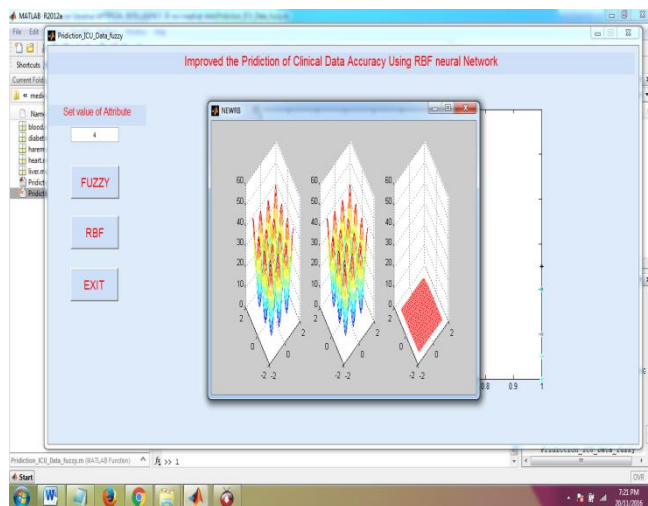
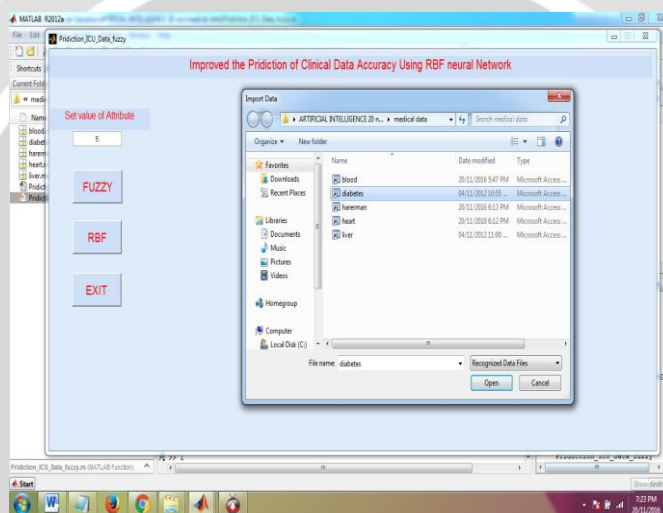**Figure 3:** Shows that the main RBF output window with value of attributes 4 and selected blood dataset.

Figure 4: Shows that the main implementation window with value of attributes 5 and selection of diabetes dataset.

METHOD	ELAPSED TIME	ACCURACY	SENSITIVITY	SPECIFICITY
FUZZY	12.2	85.9	99.0	93.8
RBF	69.6	95.9	101.0	93.3

Table 1: Shows that the comparative result analysis for input attributes value is 4 and blood dataset with using Fuzzy and RBF method and we find the value of Elapsed time, Accuracy, Sensitivity and Specificity.
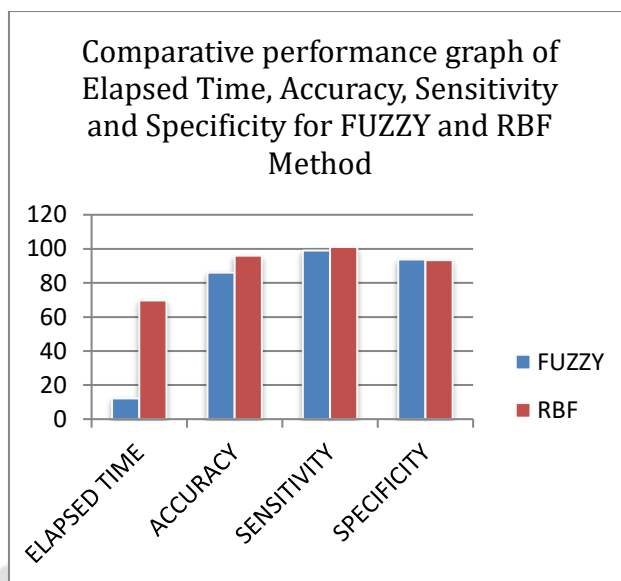
Figure 5: Shows that the comparative results graph for input attributes value is 4 and blood dataset with using Fuzzy and RBF method.
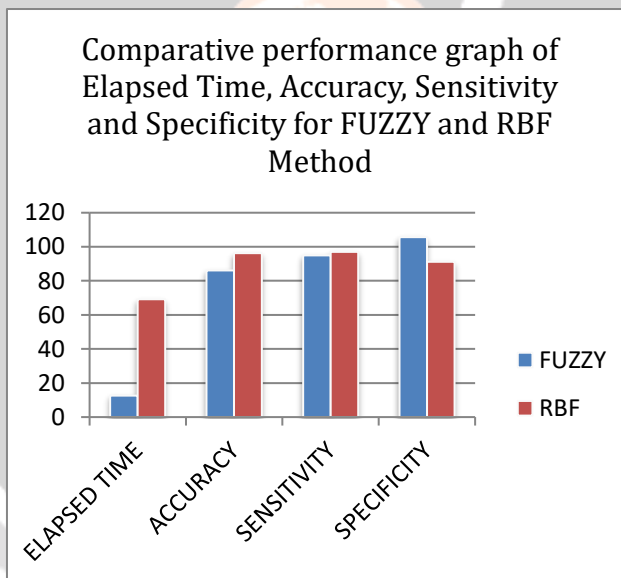

Figure 6: Shows that the comparative results graph for input attributes value is 5 and diabetes dataset with using Fuzzy and RBF method.

6.CONCLUSION AND FUTURE WORK

The information mining method is pail of calculation for the forecast and order of ICU information. The fell RBF arrange enhanced the exactness of minority class of classifier and lessens the unclassified information in ID3 characterization. The expanding of ID3 characterization area enhanced the precision and execution of classifier. Our observational result indicates better result in pressure of DT adjusted information in ID3 grouping. The fell RBF organize additionally enhanced the execution of classifier as far as unpredictability of calculation. We appeared through exploratory results that the fell calculation is compelling in the preparation of both RF and neural systems. In this paper we proposed a hybrid method for ICU data classification for ID3 classification. Our experimental result shows that better result in compression of old and traditional method of blood data classification. However, the computational time of process is increment. In future we utilized inspecting strategy for the diminishment of time and change of minority class grouping. Furthermore, another work of future is RBF testing process connected in irregular backwoods characterization for better mapping of highlight space.

REFERENCES

[1] Yi-Zeng Hsieh, Mu-Chun Su, Chen-Hsu Wang and Pa-Chun Wang "Prediction of survival of ICU patients using computational intelligence", Elsevier, 2023, Pp 13-19.

[2] F. Cismondi, L.A. Celi, A.S. Fialho, S.M. Vieira, S.R. Reti, J.M.C. Sousa and S.N. Finkelstein "Reducing unnecessary lab testing in the ICU with artificial intelligence", international journal of medicalinformatics, 2023, Pp 345-358.

[3] Thomas Berlet "RATIONALISING STANDARDLABORATORY MEASUREMENTS IN THE INTENSIVE CARE UNIT", ICU Management, 2022, Pp 33-35.

[4] Aisyah Hartini Jahidin, Mohd Nasir Taib, Nooritawati Md Tahir and Megat Syahirul Amin Megat Ali "IQ Classification via Brainwave Features: Review on Artificial Intelligence Techniques", IJECE, 2021, Pp 84-91.

[5] Rúben Duarte M. A. Pereira, Cátia M. Salgado, Andre Dejam, Shane R. Reti,Susana M. Vieira, João M. C. Sousa, Leo A. Celi and Stan N. Finkelstein "Fuzzy Modeling to Predict Severely Depressed Left Ventricular Ejection Fraction following Admission to the Intensive Care Unit Using Clinical Physiology", Scientific World Journal, 2021, Pp 1-10.

[6] Gavin Robertson, Eldon D. Lehmann, William Sandham and David Hamilton "Blood Glucose Prediction Using Artificial Neural Networks Trained with the AIDA Diabetes Simulator: A Proof-of-Concept Pilot Study", Journal of Electrical and Computer Engineering, 2021, Pp 1-12.

[7] Ashish Kumar Sen, Shamsher Bahadur Patel and Dr. D. P. Shukla "A Data Mining Technique for Prediction of Coronary Heart Disease Using Neuro-FuzzyIntegrated Approach Two Level", International Journal Of Engineering And Computer Science, 2020, Pp 2663-2671.

[8] Rosma Mohd Dom, Basir Abidin, Sameem Abdul Kareem, Siti Mazlipah Ismail and Norzaidi Mohd Daud "Determining the Critical Success Factors of Oral Cancer SusceptibilityPrediction in Malaysia Using Fuzzy Models", Sains Malaysiana, 2020, Pp 633-640.

[9] Jheng-Yan Lan, Maysam F. Abbod, Rong-Guan Yeh, Shou-Zen Fan and Jiann-Shing Shieh "Review: Intelligent Modeling and Control in Anesthesia", Journal of Medical and Biological Engineering, 2020, Pp 2393-308.

[10] Gang Wang, Jinxing Hao, Jian Ma and Lihua Huang "A new approach to intrusion detection using Artificial Neural Networks and fuzzy clustering",Expert Systems with Applications, 2019, Pp 1-8.

[11] Konstantia Zarkogianni, Andriani Vazeou, Stavroula G. Mougiakakou, Aikaterini Prountzou, and Konstantina S. Nikita "An Insulin Infusion Advisory System Based on Autotuning Nonlinear Model-Predictive Control", IEEE, 2018, Pp 2467-2477.

[12] Dr. Anooj P.K. "Prediction Of Heart Disease Using Decision Tree Fuzzy Rules", Asian Transactions on Computers, 2018, Pp 1-11.

[13] J. Chen, K. Chandrashekhara, C. Mahimkar, S.N. Lekakh and V.L. Richards "Void closure prediction in cold rolling using finite element analysis and neural network ", Journal of Materials Processing Technology, 2018, Pp 245–255.

[14] Elpiniki I. Papageorgiou "A Fuzzy Inference Map approach to cope with uncertainty in modeling medical knowledge and making decisions", Intelligent Decision Technologies, 2019, Pp 1-17.

[15] D.A. Mishra and A. Basu "Estimation of uniaxial compressive strength of rock materials by index tests using regression analysis and fuzzy inference system", Elsevier, 2018, Pp 54-68.