

Early prediction of diabetes diagnosis using hybrid classification techniques

Lakshmi Shrinivasan¹, Reshma Verma¹, Mysore Dakshinamurthy Nandeesh²

¹Department of Electronics and Communication Engg, Faculty of ECE, Ramaiah Institute of Technology, Bangalore India

²Department of Electronics and Instrumentation Engg, Faculty of EIE, Ramaiah Institute of Technology, Bangalore, India

Article Info

Article history:

Received Jan 11, 2022

Revised Jun 21, 2022

Accepted Jul 16, 2022

Keywords:

Centroid

Data mining

Decision tree

Fuzzy inference system

Modified adaptive neuro fuzzy inference system

Optimal model

Type 2 fuzzy logic

ABSTRACT

Diabetes can be mentioned as one of the most lethal and constant sicknesses that may cause an arise in the glucose levels. Design and development of performance efficient diagnosis tool is important and plays a vigorous role in initial prediction of disease and help medical experts to start with suitable treatment or medication. The insulin produced by pancreases in the subject's body will be affected leading to several dysfunctionalities to various body organs such as kidney, heart eyes and nervous system with their normal functionalities. Hence, preliminary stage detection with proper care and medication could reduce the risk of these problems. In the area of medicine to discover patient's data as well as to attain a predictive model or a set of rules, classification techniques have been continuously used. This study helped diagnose diabetes by selecting three important artificial intelligence (AI) techniques namely the optimal decision tree algorithm model, Type-2 fuzzy expert system and adaptive neuro fuzzy inference system which is modified. In the present research work, a hybrid model is proposed in order to improve the classification prediction and accuracy. The Pima Indian diabetes dataset (PIDD) from machine learning repository dataset was used to carry out validation and predication of the model accuracy.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Lakshmi Shrinivasan

Department of Electronics and Communication Engineering, Ramaiah Institute of Technolgy

Karnataka, India

Email: lakshmi15121977@rediffmail.com

1. INTRODUCTION

The knowledge based expert support systems based on decision (DSSs) are utilized in the field of medicine to support doctors and health care experts to make suitable decisions in various domains [1]. Based on knowledge and domain expert with relevant understanding of a relationship for detection, identification and diagnosis of facts in medical areas which results in analyzing of data and relevant treatment for the identified disease. In recent years artificial intelligence (AI) has acquired great weightage in researchers' community in different domains specially in medical fields for diagnosis and prediction of disease. DSS which is based on AI are more effective, accurate and reliable as compared to other support systems [1]. Now these latest years, people's life style has been resulted in occurrence of diseases. Diabetes mellitus as a kind of normal disease. In this, the body is unable to generate or respond to required level of insulin resulting in abnormal metabolism of carbohydrates, heart at risk, kidney may damage and increased amount of sugar in the blood and urine [2]. A knowledge based expert system based on AI that incorporates knowledge to resolve a complex problems and assist a human expert in making accurate decisions in medical domain [3], [4]. Researchers have gain the benefit of AI in different domains specially in medicine for diagnosis and prediction of several diseases as compared to other decision models in literature [2].

The severity levels of chronic diseases lead to various organ failure. There are two types of diabetes namely, Type 1 and Type 2. The deficiency of insulin secretion results in Type 1. Type 2 diabetes is much more dominant, this is due to resistance to insulin action and an insufficient insulin secretory response. Type 2 diabetes leads to increase in blood sugar levels beyond the normal range [5], [6]. Suitable and proper medication can control these severe conditions. By 2030, as Survey conducted by World Health Organization (WHO) depicts that 45 crores of people currently will be suffering from acute and chronic disease across globe and is predicted to increase almost double.

Fuzzy based expert systems handle uncertainties in data based on rules framed by domain experts as these are essential for preliminary stage diagnosis and prediction of diabetes. This system can be cost-effective beneficial for medical diagnosis detection system. These systems offer precise solution by handling linguistic data [7]. The adaptive neuro-fuzzy inference system (ANFIS) is the combination of fuzzy logic and neural network having the advantage of predictability and human learning-based models. The combined ANFIS with fuzzy inference system of Takagi–Sugeno type (TSK) which is established on fuzzy mathematical calculations which can resolve real time complex and crucial problems [8].

Detection and diagnosis of diabetes of Type 2 at initial stage is high in demand. In recent years, AI and decision support techniques are getting more prominence in medical diagnosis domain by their classification and prediction capability. In this paper, a hybrid diabetes Type 2, diagnosis and prediction model is proposed. For reduction of data a classifier termed J48 [9] decision tree has been proposed.

The key purpose of the paper is to search and provide a better diagnosis and prediction of diabetes by predicting the blood glucose level in an advance that is before 2 hours. Presently there are a quite a lot of other methodologies do employ on classification for the diabetes disease (DD). The planned methodology that has been adopted for the classification and prediction, on the selected feature is J48 Decision Tree algorithm [10].

One of the well-known techniques of discovering unknown patterns or prediction and classification rules is data mining decision tree is one of the DSS data identifying methods. A tree is used in decision as a one of the classification the techniques and it has three types of nodes: the root node, internal node, target node or end node [11], [12]. Wu and Mendel in [13], presented work on IT-2 fuzzy systems for the big data analytics and explained about the types of membership functions used and also about the type reducer which is an extra block used compared to type 1 fuzzy systems. Vidhya and Shanmugalakshmi in their paper [14] proposed an adaptive neuro fuzzy inference system which is modified (M-ANFIS) for various disease analysis of medical systems. Big data by calculating the closed frequent item set and their entropies.

With these relevant literature study proved that these DSS [15]–[19] systems using available dataset can predict and diagnosis diseases with least error. In the present work, a proposed hybrid model with different classification algorithms has been developed. The algorithms were the decision tree was implemented followed by design of fuzzy expert system using type 1 and Type 2 fuzzy logic and M-ANFIS using Pima Indian diabetes dataset (PIDD) which predicts the occurrence/early stages of diabetes. Later the obtained results were compared with the decision tree model. The performance metrics depicted that M-ANFIS model offers improved classification prediction and accuracy with least error.

2. METHODOLOGY

The proposed methodology is depicted as shown in Figure 1. This proposed model comprises of different AI based soft computing algorithms. It is used to predict the accuracy of diagnosing of the type 2 diabetes at early stage.

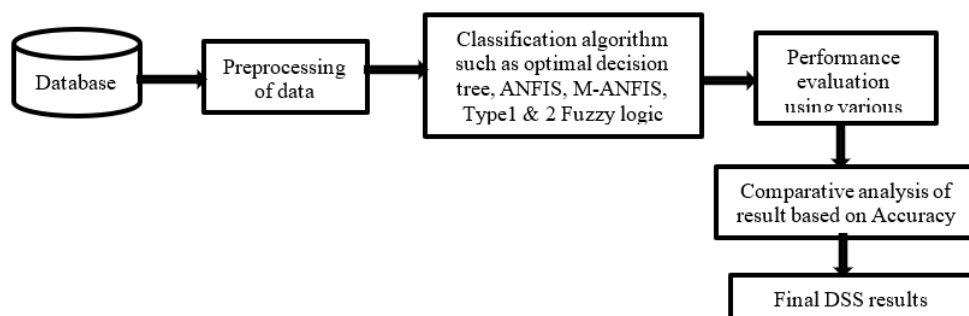


Figure 1. Proposed hybrid methodology

2.1. Optimal decision tree algorithm

In the proposed technique, a simple but efficient optimal decision tree-based classification and prediction of type 2 diabetes were presented. The algorithm was designed and developed using the PIMA dataset. In order to control level of branches of the proposed tree model, an expectation-maximization clustering method is used for minimization of data, and the datasets which further bifurcated into three different sets. Selecting various parameters, the decision tree is constructed. Later optimum tree model is selected based on the greatest accuracy of the model.

In PIDD dataset has '0' or '1' combinations which indicates negative and positive diagnosis respectively. The proposed algorithm determines the likelihood of happening of diabetes disease in a subject based on laboratory tests and signs observed. Diabetes can be observed in both genders in different age categories. The inputs specified were body mass index (BMI), glucose level, insulin level, age and diabetes pedigree function (DPF) from each subject to obtain and compute possibility of happening of diabetes.

2.1.1. Tree based on decision algorithm

Data mining is one of the well-known techniques in recognizing unknown patterns and prediction rules. A decision tree is one among many data mining techniques. This a classification technique which was represented as a tree structure, and it has mainly three kinds of nodes namely, the end or target, internal and root node. An attribute is denoted by each and every node, branch denotes attribute, and a class by a leaf node. Categorical and numerical features are effectively handled and represented by decision tree algorithm. To begin with algorithm implementation during its initial step it sets the input and output attributes. The output was represented as zero or unity i.e., likelihood of happening of diabetes based on the given input parameters. Design and implementation of decision tree was carried out using Matlab toolbox.

A. Expectation maximization clustering

Gaussian Function model is represented as each cluster. In this expectation maximization (EM) uses a given multivariate Gaussian likelihood distribution function. A given data point whether it will belong to a cluster or not based on this model. This is achieved by two interchanging steps.

- a) Expectation (E-Step): The weight of the data point is calculated based on the likelihood of it belonging to each cluster. The chances of occurring a given a point is more, if a point is belonging to a cluster, then it's been assigned to a very close value to 1. Probability distribution of clustering of data points is established for the situation where a point may fit to two or more groups.
- b) Maximization M-Step: Using the weight calculated for each point in previous step, it is used to approximate the relevant parameters of each group. In E-Step every data point is weighted with likelihood, and the variance and mean of each group were obtained. The maximum complete likelihood clustering is found. Until the convergence occurs, the Expectation and maximization steps will be repeated constantly to enhance the entire likelihood which is logarithmic. To prevent local optimization Multiple iterative steps are required. The Decision tree selection process is controlled by the various hyper parameters are extreme bins, extreme depth, contamination measure, and least info gain. Extreme depth which bounds the number of levels in this algorithm process. To classify a given sample, a classifier makes a number of continuous decisions for a given data sample points. To avoid overfitting, the number of stages is limited in tree.

B. Classification process

The steps for the classification and prediction of diabetes using decision tree are showed in Figure 2. First step is data preprocessing uses EM clustering algorithm, not correctly classified data was eliminated. The data is then classified into different data sets for model evaluation, testing and validation.

Randomly selected the 70% data is treated as training dataset as 1 and remaining 30% of the dataset is used to test and evaluate the performance of the DSS tree algorithm. It is easy to over fit the model as the model has been tested many times. Further to gain more accuracy and validation of training model dataset 1 is bifurcated into two subsets as 90% and 10% dataset. For training 90% dataset is utilized which is called set 2 training. The cross-validation (CV) is carried out with 10% leftover dataset to evaluate the model. The total dataset is classified into three subsets, namely set 2 training, the set CV, and set test as shown in Figure 3.

Training model identify the suitable hyper parameter for the decision tree algorithm. A model needs to be built and tested for the hyper parameter combination. The decision tree model is trained using 70% of the data set. For prediction optimal decision tree model will be selected after training the model with the dataset. Depending on the probability of occurrence values the prediction of diabetes is calculated.

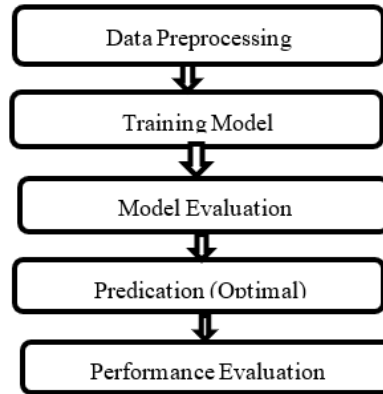


Figure 2. Proposed classification DSS model for diabetes

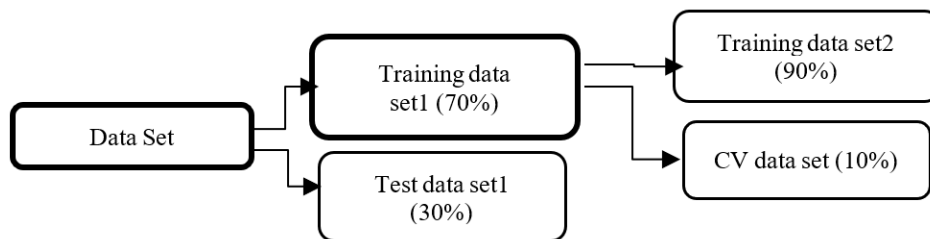


Figure 3. Partitioned datasets

C. Model evaluation and performance evaluation

The accuracy and predicting capability of a classifier is effectively calculated. For predicting the class tag as a dataset of tuples will be mentioned. Mainly three parameters were focused they are as accuracy, sensitivity and specificity: Using Confusing Matrix four evaluation metrics were calculated.

- True positive (TP): Correctly tagged positive tuples by a classifier.
- True negative (TN): The negative tuples that were correctly tagged by the classifier.
- False positive (FP): Incorrectly tagged negative tuples as positive.
- False negative (FN): The positive tuples that were wrongly tagged as negative. In proposed work, accuracy, sensitivity and specificity were calculated using (1)-(3).

$$Accuracy = \frac{TP+FN}{TP+TN+FP+FN} \quad (1)$$

$$Specificity = \frac{TN}{TN+FP} \quad (2)$$

$$Sensitivity = \frac{TP}{TP+FN} \quad (3)$$

2.2. Interval type-2 fuzzy expert system

Fuzzy system is one of the most efficient qualitative computational method which can manage large ambiguous dataset to provide precise results. Here, the system variables are defined as linguistic terms and fuzzy rules are generated to model the imprecise aspects of system behavior. Fuzzy logic [20] is useful more so for its easy implementation and speedy generation of results.

The control decision based support system (CDSS) tool incorporated with fuzzy logic is referred to as a knowledge-based system that contains both static and dynamic information. Considering the large amount of diagnosis data which is ambiguous in nature, the analysis of these qualitative and quantitative variables has to be carried out efficiently to take appropriate decision. In the proposed DSS system, initially the input and output features were defined. The fuzzy output was labeled as very low (VL), low (L), medium low (ML), high (H), very high (VH) based on diagnostic likelihood. The block diagram for type-2 fuzzy logic system is shown in Figure 4.

Defuzzification is a process in which centroid method is employed and it's the reduction of type 2 to type 1 fuzzy logic crisp output. In this paper, type reduction is briefly described. The centroid type-reduction is illustrated Fuzzification is carried out using Triangular membership function using knowledge base rules as illustrated in Figure 5 and membership function forage attribute is shown in Figure 6. The single crisp output value is obtained by defuzzification using centroid method. The crisp output is the result of the probability of diagnosis. The rule base generated in MATLAB is shown in Figure 7.

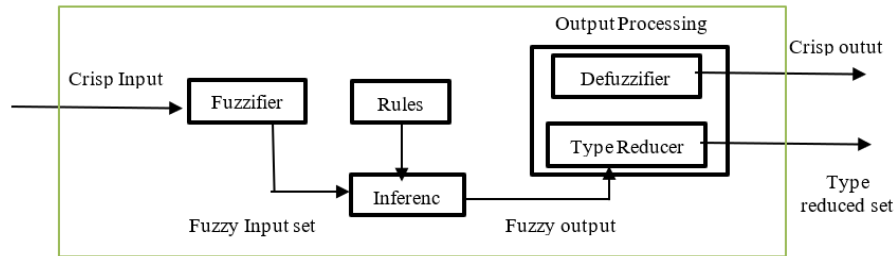


Figure 4. Type-2 fuzzy logic system

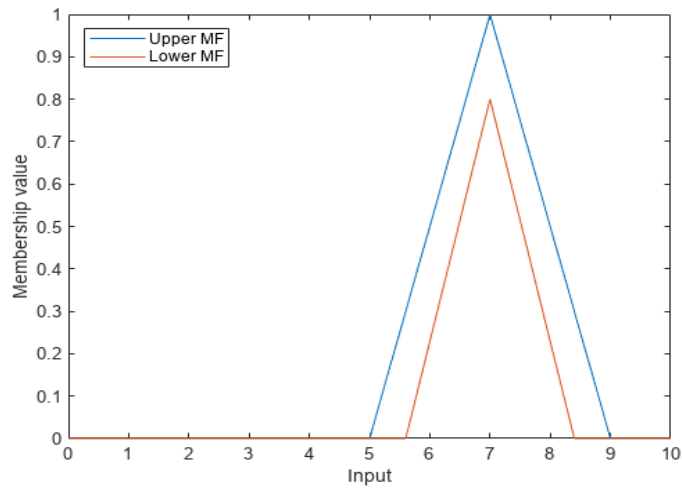


Figure 5. Triangular membership function for type 2 fuzzy

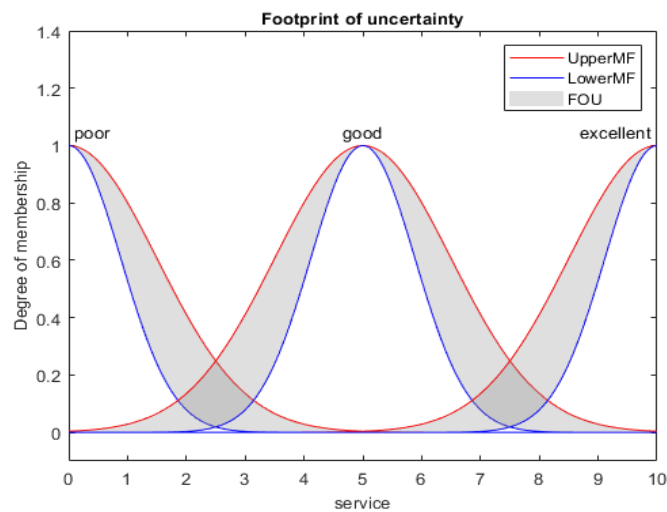


Figure 6. 3D membership function for service which is age input attribute using type2 fuzzy

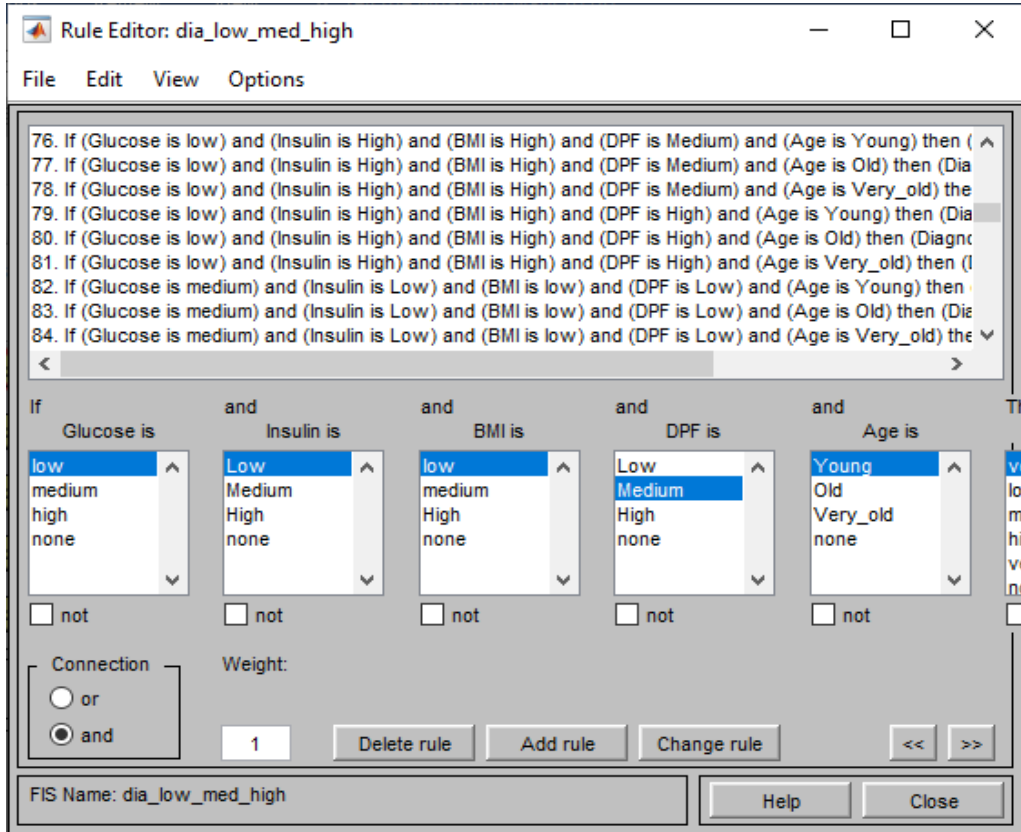


Figure 7. Domain experts based 240 rules generated in MATLAB

2.3. Adaptive neuro fuzzy inference system-modified (M-ANFIS)

The fuzzy expert system with features such as powerful decision support, deduction of consequences which is based on the uncertain information stored in knowledge base and enhanced reasoning capability is used as a decision-making tool in critical situations. But fuzzy expert system lacks the ability to adapt to changing environment. Neural network [21]–[25], on the other hand, has the ability to learn, process information in parallel to classifying statistical parameters. In this approach, the knowledge accumulates in the form of weights in between the neurons (nodes) at the connections, which is adjusted in the learning process to adapt to the changing input.

The data prediction classification and analysis of diabetes disease was carried out using M-ANFIS. Initially, the PIMA dataset undergoes pre-processing. Data formatting, identifying has been carried out in pre-processing stage on diabetes dataset. then, feature extraction was done and the count frequent item (CFI) for the closed is obtained. Afterwards, using CFI count the entropy is determined.

Figure 8 shows the structure of the proposed ANFIS DSS model which is made up of three adjustable nodes and two fixed nodes those are connected with first order fuzzy system which is of TSK model. It is made up of five input features and one output. Fuzzy inference system (FIS) was generated with triangular membership function. The models were optimized and classified using hybrid algorithm. Further, classification results were compared with that obtained by applying back propagation algorithm. A total of 3-3-3-3-3 numbers of membership functions were selected corresponding to inputs glucose, serum insulin level, BMI, DPF and age respectively. 3-3-3-3-3=243 if-then fuzzy rules were created such that T-norm operators connected fuzzy parameters using fuzzy and operators. The testing and training process were carried out and the performance metric for root mean square error (RMSE) were obtained using equation (4) for the designed and developed model, where total number of input data A, ti is ith measured value and qi is its prediction value. Both training and testing process were repeated for different epoch values as shown in Figure 8.

$$RMSE = \sqrt{\frac{\sum_{i=1}^A [t_i^2 - q_i^2]}{A}} \tag{4}$$

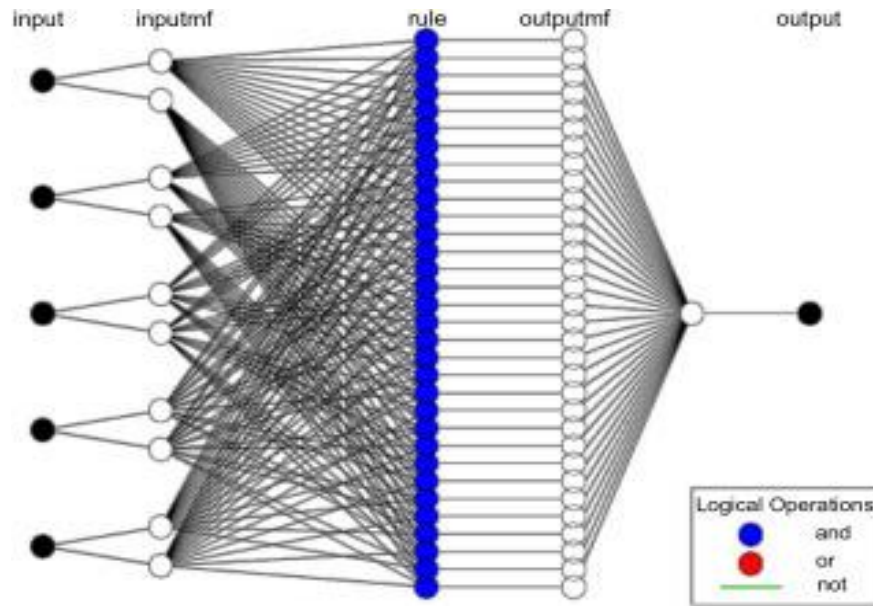


Figure 8. Proposed ANFIS DSS model

3. RESULTS AND DISCUSSION

In the optimal decision tree algorithm, incorrectly sample classification was removed by means of EM clustering. After the entire number of samples were divided into training and test set. Decision tree model was trained by using the training set data. The graph of probability of occurrence of diabetes against glucose values is shown in Figure 9. The regression tree is shown in Figure 10. Accuracy, sensitivity parameters are calculated by finding the confusion matrix.

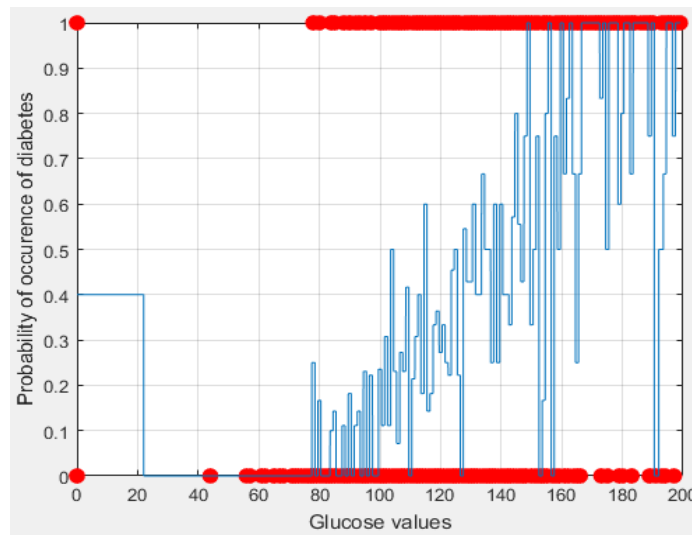


Figure 9. The likelihood of happening of diabetes VS glucose values

After the decision tree algorithm fuzzy models were generated the output of the fuzzy set system model were divided into five different groups namely, very low (VL), low (L), medium low (ML), medium high (MH), very high (VH) based on the likelihood or severity of diabetes. Using matlab fuzzy tol box the outcome of probability of diabetes for the two distinct cases, one with chances of low probability based on age and age is obtained as 0.208 and 0.66 for chance of high probability outcomes of the system. Figure 11 provides 3D surface plot view of the impact of the inputs glucose and age used in diabetes diagnosis.

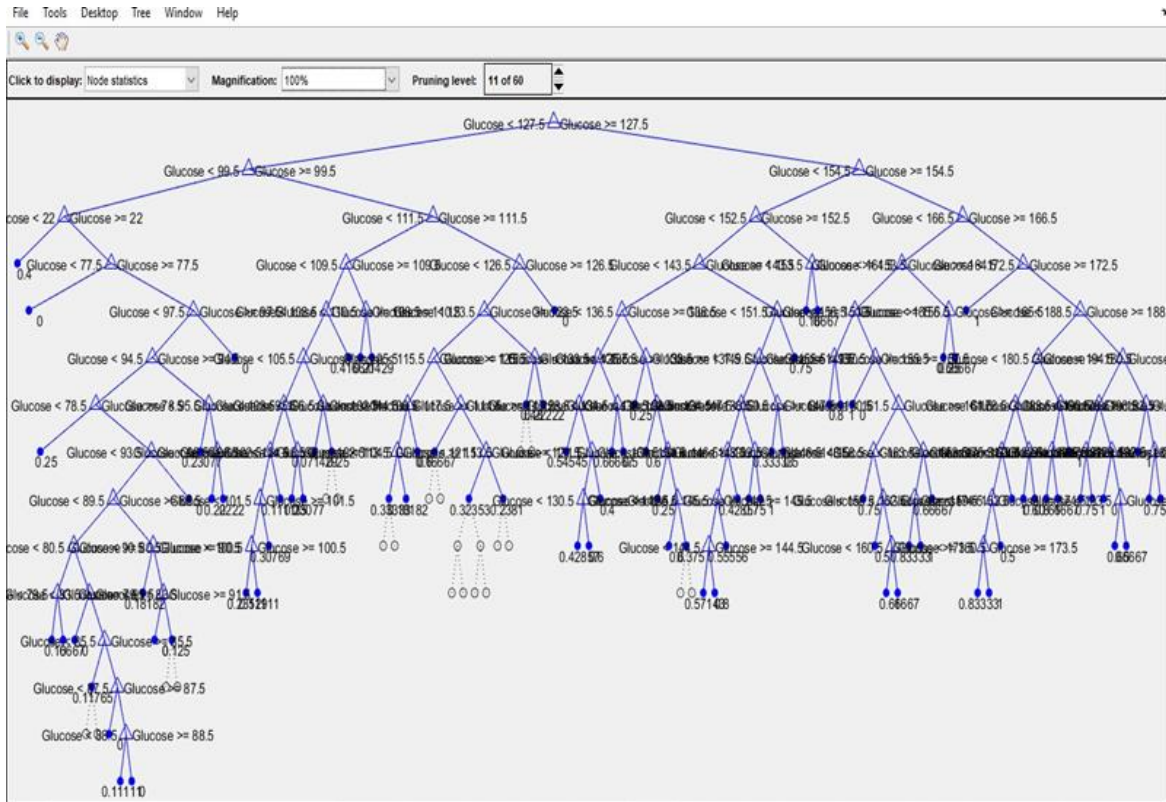


Figure 10. Regression tree

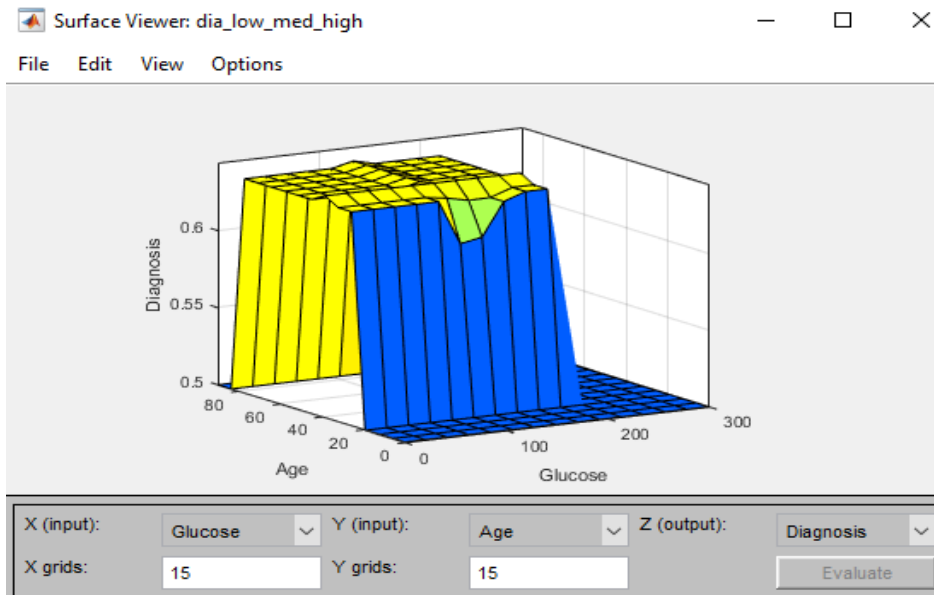


Figure 11. 3D surface plot depicting the impact of inputs glucose and age on the diagnosis

Later, ANFIS model was design and developed using simulation fuzzy logic toolbox which was trained, tested and validated. The model was validated for performance efficiency and accuracy using RMSE metrics for different EPOCHs and depicted in Table 1 and performance validation of different proposed model are listed in Table 2 also RMSE and MSE performance metrics are depicted in Table 3. All the statistical parameters were generated using confusion matrix.

Table 1. Hyrid algoritmt

	During training and testing				
	10	20	50	100	150
RMSE	0.0623	0.0623	0.0623	0.0623	0.0626
MSE	0.004386	0.004386	0.004386	0.004386	0.004386

Table 2. Statistical performance parameter metrics of different classification DSS

Different classification techniques	Accuracy (%)	Sensitivity (%)	Specificity (%)
ANFIS [1]	86.48	92.78	73.08
Optimal Decision Tree Algorithm	92.8	93.2	92.1
M-ANFIS	97.5	96.9	95.6
KNN	77.1	79.32	70.40

Table 3. Comparison of performance validation of different classification methods in terms of RMSE and

Classification models	MSE values	
	RMSE	MSE
Type 1 Fuzzy Logic	0.45924	0.21090
Type 2 Fuzzy Logic	0.2283	0.05212
ANFIS	0.21964	0.04824
M-ANFIS	0.06623	0.004386

The results from Table 2 and Table 3 proved that M-ANFIS classification model had least errors by observing RMSE and MSE performance metrics values. The accuracy has been improved in comparison with other classification diagnosis techniques and algorithms. Hence the M –ANFIS proposed hybrid model shows better performance. The accuracy of model is 97.5 which is better when compared to other algorithms for early detection of diabetes.

4. CONCLUSION

Precise, accurate, robust and reliable DSS systems are essential in medical field. The results obtained which are mentioned in Table 2 and Table 3 showed that M-ANFIS classification model had better accuracy and least error as compared with other classification techniques. Comparing with the results, it is proved that the proposed hybrid model is more efficient in classifying the possible type 2 diabetes at early stages for subjects. The proposed model is validated with PIDD data only. Further, from the practical implementation of the model it needs to be trained and validated with different large number of datasets to assess the robustness of the model also hardware prototype model as a device could be implemented.




REFERENCES

- [1] A. Yahyaoui, A. Jamil, J. Rasheed, and M. Yesiltepe, "A decision support system for Diabetes prediction using machine learning and deep learning techniques," *1st International Informatics and Software Engineering Conference: Innovative Technologies for Digital Transformation, IISEC 2019 - Proceedings*, 2019, doi: 10.1109/UBMYK48245.2019.8965556.
- [2] W. Kerner and J. Brückel, "Definition, classification and diagnosis of diabetes mellitus," *Experimental and Clinical Endocrinology and Diabetes*, vol. 122, no. 7, pp. 384–386, 2014, doi: 10.1055/s-0034-1366278.
- [3] P. S. K. Patra, D. P. Sahu, and I. Mandal, "An expert system for diagnosis Of human diseases," *International Journal of Computer Applications*, vol. 1, no. 13, pp. 71–74, 2010, doi: 10.5120/279-439.
- [4] C. Yau and A. Sattar, "Developing expert system with soft systems concept," *Proceedings of the IEEE International Conference on Expert Systems for Development*, pp. 79–84, 1994, doi: 10.1109/icesd.1994.302302.
- [5] C. D. Malchoff, "Diagnosis and classification of diabetes mellitus," *Diabetes Care*, vol. 37, no. SUPPL.1, 2014, doi: 10.2337/dc14-S081.
- [6] O. Geman, I. Chiuchisan, and R. Todorean, "Application of adaptive neuro-fuzzy inference system for Diabetes classification and prediction," *2017 E-Health and Bioengineering Conference, EHB 2017*, pp. 639–642, 2017, doi: 10.1109/EHB.2017.7995505.
- [7] B. Pandey and R. B. Mishra, "Knowledge and intelligent computing system in medicine," *Computers in Biology and Medicine*, vol. 39, no. 3, pp. 215–230, 2009, doi: 10.1016/j.compbiomed.2008.12.008.
- [8] R. Seising, "From vagueness in medical thought to the foundations of fuzzy reasoning in medical diagnosis," *Artificial Intelligence in Medicine*, vol. 38, no. 3, pp. 237–256, 2006, doi: 10.1016/j.artmed.2006.06.004.
- [9] W. Chen, S. Chen, H. Zhang, and T. Wu, "A hybrid prediction model for type 2 diabetes using K-means and decision tree," *Proceedings of the IEEE International Conference on Software Engineering and Service Sciences, ICSESS*, vol. 2017-Novem, pp. 386–390, 2018, doi: 10.1109/ICSESS.2017.8342938.
- [10] K. R. Pradeep and N. C. Naveen, "Predictive analysis of diabetes using J48 algorithm of classification techniques," *Proceedings of the 2016 2nd International Conference on Contemporary Computing and Informatics, IC3I 2016*, pp. 347–352, 2016, doi: 10.1109/IC3I.2016.7917987.
- [11] H. Zhang and R. Zhou, "The analysis and optimization of decision tree based on ID3 algorithm," *Proceedings of 2017 9th International Conference On Modelling, Identification and Control, ICMIC 2017*, vol. 2018-March, pp. 924–928, 2018,




- doi: 10.1109/ICMIC.2017.8321588.
- [12] Z. Sun, S. Yu, and Y. Zhang, "An optimal decision tree model for diabetes diagnosis," *Proceedings - 2019 4th International Conference on Computational Intelligence and Applications, ICCIA 2019*, pp. 83–87, 2019, doi: 10.1109/ICCIA.2019.00023.
- [13] D. Wu and J. M. Mendel, "Recommendations on designing practical interval type-2 fuzzy systems," *Engineering Applications of Artificial Intelligence*, vol. 85, pp. 182–193, 2019, doi: 10.1016/j.engappai.2019.06.012.
- [14] K. Vidhya and R. Shanmugalakshmi, "Modified adaptive neuro-fuzzy inference system (M-ANFIS) based multi-disease analysis of healthcare big data," *Journal of Supercomputing*, vol. 76, no. 11, pp. 8657–8678, 2020, doi: 10.1007/s11227-019-03132-w.
- [15] L. Shrinivasan and J. R. Raol, "Interval type-2 fuzzy-logic-based decision fusion system for air-lane monitoring," *IET Intelligent Transport Systems*, vol. 12, no. 8, pp. 860–867, 2018, doi: 10.1049/iet-its.2017.0095.
- [16] W. Yu, T. Liu, R. Valdez, M. Gwinn, and M. J. Khoury, "Application of support vector machine modeling for prediction of common diseases: The case of diabetes and pre-diabetes," *BMC Medical Informatics and Decision Making*, vol. 10, no. 1, 2010, doi: 10.1186/1472-6947-10-16.
- [17] K. M. Orabi, Y. M. Kamal, and T. M. Rabah, "Early predictive system for diabetes mellitus disease," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9728, pp. 420–427, 2016, doi: 10.1007/978-3-319-41561-1_31.
- [18] A. S. Rawat, A. Rana, A. Kumar, and A. Bagwari, "Application of multi layer artificial neural network in the diagnosis system: A systematic review," *IAES International Journal of Artificial Intelligence*, vol. 7, no. 3, pp. 138–142, 2018, doi: 10.11591/ijai.v7.i3.pp138-142.
- [19] I. Kavakiotis, O. Tsave, A. Salifoglou, N. Maglaveras, I. Vlahavas, and I. Chouvarda, "Machine learning and data mining methods in Diabetes research," *Computational and Structural Biotechnology Journal*, vol. 15, pp. 104–116, 2017, doi: 10.1016/j.csbj.2016.12.005.
- [20] K. M. Aamir, L. Sarfraz, M. Ramzan, M. Bilal, J. Shafi, and M. Attique, "A fuzzy rule-based system for classification of diabetes," *Sensors*, vol. 21, no. 23, 2021, doi: 10.3390/s21238095.
- [21] Y. Zhang, Z. Lin, Y. Kang, R. Ning, and Y. Meng, "A feed-forward neural network model for the accurate prediction of diabetes mellitus," *International Journal of Scientific and Technology Research*, vol. 7, no. 8, pp. 151–155, 2018.
- [22] M. Ayad, H. Kanaan, and M. Ayache, "Diabetes disease prediction using artificial intelligence," *Proceedings - 2020 21st International Arab Conference on Information Technology, ACIT 2020*, 2020, doi: 10.1109/ACIT50332.2020.9300066.
- [23] E. M. Marouane and Z. Elhoussaine, "A fuzzy neighborhood rough set method for anomaly detection in large scale data," *IAES International Journal of Artificial Intelligence*, vol. 9, no. 1, pp. 1–10, 2020, doi: 10.11591/ijai.v9.i1.pp1-10.
- [24] S. Perveen, M. Shahbaz, K. Keshavjee, and A. Guergachi, "Metabolic syndrome and development of Diabetes Mellitus: Predictive modeling based on machine learning techniques," *IEEE Access*, vol. 7, pp. 1365–1375, 2019, doi: 10.1109/ACCESS.2018.2884249.
- [25] A. Iyer, J. S, and R. Sumbaly, "Diagnosis of Diabetes using classification mining techniques," *International Journal of Data Mining & Knowledge Management Process*, vol. 5, no. 1, pp. 01–14, 2015, doi: 10.5121/ijdkp.2015.5101.

BIOGRAPHIES OF AUTHORS






Dr Lakshmi Shrinivasan    holds a PhD degree from Jain University, India in 2018. She also received her B.E. from Maharashtra and M.Tech. (Electronics) from VTU, Karnataka, India in 1999 and 2007, respectively. She is currently working as an Associate professor in Dept of ECE at Ramaiah Institute of Technology, Bangalore. Her research interests are Artificial Intelligence, Embedded system Design, IoT and Robotics. She had published several research papers in International, National conferences and indexed journals. She has 15 years of academic teaching and 2 years of Industry experience. She is member of several professional bodies like MIEEE, Fellow MIETE, LMISTE and MIAENG. She has guided several PG and UG projects. She can be contacted at email: lakshmi15121977@rediffmail.com.



Dr. Reshma Verma    received the Ph.D degree from Jain University, India in 2018. Currently she is working as an Assistant Professor in the Department of ECE, Ramaiah Institute of Technology, Karnataka, India. Her research interest is Artificial Intelligence, Image processing and Embedded system design. She has published several National, international conferences and indexed journals. She is keen to work in Biomedical signal processing. She can be contacted at email: reshmaverma@msrit.edu.



Dr Nandeesh M D    received Bachelor Degree in Instrumentation Technology from Mysore University, SJCE Mysore in 1997, and M Tech degree in Biomedical Instrumentation form Visvesvaraya Technological University, Belgaum, Karnataka state, India in year 2000. Completed his Ph.D. from form Visvesvaraya Technological University, Belgaum, Karnataka state in year 2020. He is currently working as Assistant Professor, Department of Electronics and Instrumentation Engineering, M S Ramaiah Institute of Technology, Bangalore. His area of interest is Image processing, Biomedical Signal processing and pattern recognition. He has presented many papers in national and international conferences and published papers in international journals. He is a Life member of MISTE. He can be contacted at email: mndandeesh@msrit.edu.