Tech Science Press

An Intelligent Medical Expert System Using Temporal Fuzzy Rules and Neural Classifier

Praveen Talari[1,*], A. Suresh[2] and M. G. Kavitha[3]

[1]Vel Tech High Tech Dr. Rangarajan Dr. Sakunthala Engineering College, Chennai, 600062, India
[2]AMET University, Chennai, 600062, India
[3]University College of Engineering, Pattukkottai, 600062, India
*Corresponding Author: Praveen Talari. Email: praveent738@gmail.com

Abstract: As per World Health Organization report which was released in the year of 2019, Diabetes claimed the lives of approximately 1.5 million individuals globally in 2019 and around 450 million people are affected by diabetes all over the world. Hence it is inferred that diabetes is rampant across the world with the majority of the world population being affected by it. Among the diabetics, it can be observed that a large number of people had failed to identify their disease in the initial stage itself and hence the disease level moved from Type-1 to Type-2. To avoid this situation, we propose a new fuzzy logic based neural classifier for early detection of diabetes. A set of new neuro-fuzzy rules is introduced with time constraints that are applied for the first level classification. These levels are further refined by using the Fuzzy Cognitive Maps (FCM) with time intervals for making the final decision over the classification process. The main objective of this proposed model is to detect the diabetes level based on the time. Also, the set of neuro-fuzzy rules are used for selecting the most contributing values over the decision-making process in diabetes prediction. The proposed model proved its efficiency in performance after experiments conducted not only from the repository but also by using the standard diabetic detection models that are available in the market.

Keywords: Diabetes; type-1; type-2; feature selection; classification; fuzzy rules; fuzzy cognitive maps; classifier

1 Introduction

Diabetes is one of the deadliest diseases today which affects the body either by not producing sufficient insulin by pancreas or the body succumbs to the position so as to not use the insulin produced properly resulting in increased the blood glucose levels that may damage the heart, blood vessels, eyes, kidneys or the nerves. The Type-1 diabetes is classified by the shortage of insulin production in the human body. The people with Type-1 diabetes need to monitor the insulin level for regulating the glucose level in their blood. The Type-1 diabetics will not able to survive when they fail to access the insulin. Neither there is enough awareness about seriousness of Type-1 diabetes nor the measures to prevent it. The symptoms of Type-1 include the excessive urination, continuous hunger pangs, abnormal weight reduction, vision

issues and fatigue. On the other hand, the Type-2 diabetes affects the human body due to inefficient use of insulin. The Type 2 level diabetes affects majority of people in the world. The symptoms of Type-2 diabetes are similar to the Type-1 diabetes and it becomes a lifelong issue.

Feature Selection is an important part in data pre-processing method in data mining which is used to reduce the dimensionality of the given input data of dataset. In healthcare applications, Feature Selection is vital for identifying the most significant risks. This identification process is used to remove the useless as well as redundant features from the input dataset. Known to be variable selection in data pre-processing method in data mining, this process is basically utilized for reducing the data size by removing the irrelevant features from datasets. In addition, this method improves the data clarity and visualization, reduces the training time and enhances the prediction of the classifier over the datasets. Before applying any feature selection method, the removal of noisy and inconsistent data is necessary for achieving better classification accuracy with minimal time consumption. In healthcare area, many feature selection methods such as filter method, wrapper method and ensemble method are used for identifying relevant features. The dimensionality reduction is most important for the today's real world applications.

Temporal reasoning consists of four important tasks such as analyzing the diabetic's past data, predicting the effects of the prescribed medicines, identifying the modifications required, duration of the medication and the severity of the issue. Fuzzy set is useful for making decision over the crucial and large volume of input data over the serious healthcare applications that are available in the market such as diabetic disease, heart disease, etc. due to the availability of data with uncertainty (Zadeh 1988). Uncertainty over the dataset is on the rise as inconsistency, imprecision, vagueness, non-specificity and inconsistency prevail. The Fuzzy set exploits these uncertainties as challenging issues to be tackled. In addition, they create a powerful method to handle not only with the noisy, incomplete or imprecise data, these are also useful in the process of developing the uncertain models of data which provides smarter performances than the other existing systems.

Problem Statement: The available disease prediction systems are not fulfilling the current requirement in terms of prediction accuracy. Today, early prediction is playing a major role to detect and prevent the deadly diseases. Moreover, after diagnosis the diseases also need to be measured for the severity level. It is necessary develop a disease prediction system that is capable of predicting the diseases in earlier stage and also able to measure the disease level. For this purpose, a new disease prediction system is proposed in this work for fulfilling the current requirements.

In this work, a new weighted genetic based feature selection method has been proposed which identifies and selects a set of features for predicting the TYPE 1 diabetics and a new and different set of features for predicting the TYPE 2 diabetics. The major advantages of the proposed feature selection algorithm are to select the suitable features, perform tuple reduction and to clean the input data. Moreover, a new Neuro-Fuzzy Rule and Time based Classification Algorithm with Time constraints (NFRTCA) is proposed for predicting the disease level which applies the fuzzy rules for performing the initial level classification. Groups are further refined by developing fuzzy cognitive maps that constrained with time intervals for making final data classification. This Fuzzy Cognitive Map (FCM) uses 80% of the dataset as training dataset and the remaining 20% as testing dataset. The major benefit of this newly proposed system is that, it uses fuzzy rules with time constraints for performing feature selection process, classification process and the time constraints analysis based decision making. Finally, all these decisions are given as input to a filter which work collaboratively to provide effective recommendations.

2 Literature Survey

There are many works that have been carried out by many authors who worked in the areas of medical expert systems, risk prediction systems, medical data classification and feature selection in the past. Among them, Li et al. [1] introduced a new feature selection approach which works on wrapper method by using

the standard Support Vector Machine for selecting the most contributed features over the intrusion data set. Ganapathy et al. [2] proposed a new medical expert system that uses neural networks, time constraints and fuzzy rules. In their work, they have developed fuzzy rules with the consideration of time taken for making effective decision over the diabetic datasets. Nowicki et al. [3] proposed a new solution for resolving the problem of imprecise input data over the classification by using fuzzy rough sets. Their method demonstrated fuzzification as a process which is dispensable. Both the approaches are applied in neuro–fuzzy classification algorithms and also extended to be used with logical and conjunction type of fuzzy inference. In their method, the rough fuzzification embeds with logical and conjunction type of fuzzy systems. In addition, fuzzification considered neuro–fuzzy–rough sets. Pota et al. [4] developed a new method to construct a new rule based fuzzy classification algorithm based on the features of performance with interpretability that are needed by the specific application that was used with any type of data and was specifically most useful for the medical field needs. This method is in accordance to the Naïve Bayes approximation. Therefore, the parameters based optimization process was carried out only once and separately for individual attribute, thus the computational time is very less. All these works are incorporated in the fuzzy logic for making final decision on the prediction process. Even though, these work have not attained the current requirements due to the failure of effective fuzzy rule generation.

Mohammed et al. [5] proposed a new neural network named as improved Fuzzy Min-Max (FMM) neural network with a k-Nearest Neighbour (k-NN) Hyper-box expansion rule. The major objective of their work is to reduce the FMM neural network complexity to undertake the classification. They have rectified the issues of FMM and improved the classification accuracy. Moreover, they have used K-nearest hyper-box aware rules that are generated for reducing the complexity to avoid many hyper-boxes. Finally, they have evaluated their model and proved it to be better by using standard datasets through experiments. Singh et al. [6] introduced a new and an effective attribute selection for performing effective classification over the various fields including medical, e-marketing etc. Their newly proposed system determines the linguistic variables for each attribute and also in generates different variables those are linguistic in nature for classification by using rules. They have used six different datasets to evaluate the proposed model. The drawbacks of these works are the usage of old classifiers and not including the severity level of the disease.

Kavakiotis et al. [7] conducted a brief survey of the applications that are related to the machine learning, data mining and the related tools in the research area of diabetes like medical diagnosis, disease prediction, diabetic complications, genetic background and healthcare. They have covered a wide range of machine learning approaches in which 85% are supervised learning methods and 15% of unsupervised learning approaches with association rules. Finally, they have identified Support Vector Machine (SVM) which is a used algorithm due to the success over the various datasets. Sasikumar et al. [8] proposed a new healthcare recommendation system that uses type-2 fuzzy ontology with Internet of Things (IoT) devices which is used for monitoring the human body during diets with specific drugs and foods. Their system extracts the patient's risk factors and also determines the patient's health level through wearable sensors and then recommends the suitable prescriptions for taking smart medicine and diet. Finally, they have achieved better prediction accuracy than other existing recommendation systems.

Kusy [9] developed a new neural network called Probabilistic Neural Network (PNN) that is sensitive over the cardinality of the input dataset for effective medical data diagnosis. Here, the PNN uses three types of classification methods including plugin, reinforcement learning and the conjugate gradients. Cordelli et al. [10] proposed a novel medical expert system that expresses the differentiation between the healthy subjects of Type-1 diabetics affected patients and the complications adopted in Type-1 diabetics. Cheruku et al. [11] developed a new hybrid method which combines the bat optimization with rough set theory named as rough set theory based bat miner. Their method contains two important stages. The duplicate features have been removed in the first stage and fired the fuzzy rules by reducing the proposed activation function. They have achieved highest prediction accuracy when compared with the rule based prediction systems that are

available in the literature. Umamageswari et al. [12] improved the performance of an adaptive neuro-fuzzy inference system using extreme learning machines concept with particle swarm optimization for performing effective medical data classification. Cui et al. [13] proposed a new hybrid feature selection algorithm for selecting useful features from medical datasets. Moreover, they used a new classifier which combines the improved support vector machine and genetic algorithm for tuning the sensitive feature of the method. de Carvalho et al. [14] developed a new clustering algorithm called Gaussian Kernel C-Means hard clustering which facilitates the automation for performing computation process. Their algorithm selects the important variables that are useful for performing clustering process. They have tested their system with synthetic data sets and University of California, Irvine (UCI) [15] machine learning repository data sets and achieved better classification accuracy. Kanimozhi et al. [16] developed a new disease prediction system that incorporates the neural network, fuzzy logic with temporal constraints for predicting the diseases. The drawbacks of all these works are the usage of different old neural classifier and the fuzzy rules that are not considered the diabetes type symptoms. Hence they failed to categorize the diabetes types as type-1 or type-2 and also failed to find the severity level.

Nawaz et al. [17] proposed an intelligent disease prediction system that applies optimization technique that incorporates gradient descendent optimization. They have proved that their work is better than other disease prediction systems that are available in the literature. Ali et al. [18] developed a new heart disease prediction system by applying the supervised machine learning methods that are available as standard classifiers. They have proposed that the prediction system is better in terms of prediction accuracy. Quteishat et al. [19] developed a fuzzy min-max neural classifier with genetic algorithm for predicting the diseases effectively. Sethukkarasi et al. [20] proposed a fuzzy temporal cognitive map to predict the diseases effectively and also proved as efficient as other methods that are available in this direction. Wu et al. [21] proposed a new binary logistic regression to predict the diabetes diseases with the help of their newly proposed feature selection. They have achieved better prediction accuracy, even though, their model failed to predict the severity level of the disease. Acharya et al. [22] developed a nonlinear disease prediction model that considers the blood glucose and predict the type-1 diabetes. They have predicted the uncertainty of the disease levels effectively. Ahmed et al. [23] developed a new fused machine learning method to predict the diabetes diseases that applied the Support Vector Machine (SVM) and Artificial Neural Network (ANN). They have analyzed the standard dataset for determining the diabetes disease as positive and negative. All the available works achieved better accuracy than the other works that were available in this direction. Even though, no systems are available in this direction for fulfilling the current requirements in these medical fields, this work proposes a new disease prediction system for predicting the diseases effectively by introducing new neural classifier with the incorporation of genetic algorithm, neural classifier, fuzzy logic and temporal logic.

3 System Architecture

The proposed disease prediction model is demonstrated in Fig. 1 which contains eleven major components such as University of California-Irvine (UCI) Repository Dataset, Data collection Agent, User Interface Module, Decision Manager, Spatial Agent, Fuzzy Inference, Temporal Agent, Data pre-processing phase, Classification Phase, Rule manager and Rule base.

The University of California, Irvine (UCI) Repository Dataset contains the standard datasets including diabetes datasets. The Data Collection Agent is used to collect the necessary and meaningful data from the dataset without redundancy. The User Interface Module collects the most useful feature based on the given rules and regulation from University of California, Irvine (UCI) Repository through data collection agent. The Decision Manager has the overall control over all the components that are available in the overall system architecture. The Spatial Agent is used to provide the spatial related information which is relevant for the diabetes. The Fuzzy Inference system is responsible for performing the fuzzification and de-fuzzification

processes. The Temporal Agent is used to provide the temporal information to the decision manager for making effective decision on time over the medical datasets. The Data pre-processing phase consists of two sub components such as genetic algorithm and temporal rules. Here, the weighted genetic algorithm performs data pre-processing and selects the most contributed features by applying temporal fuzzy rules.
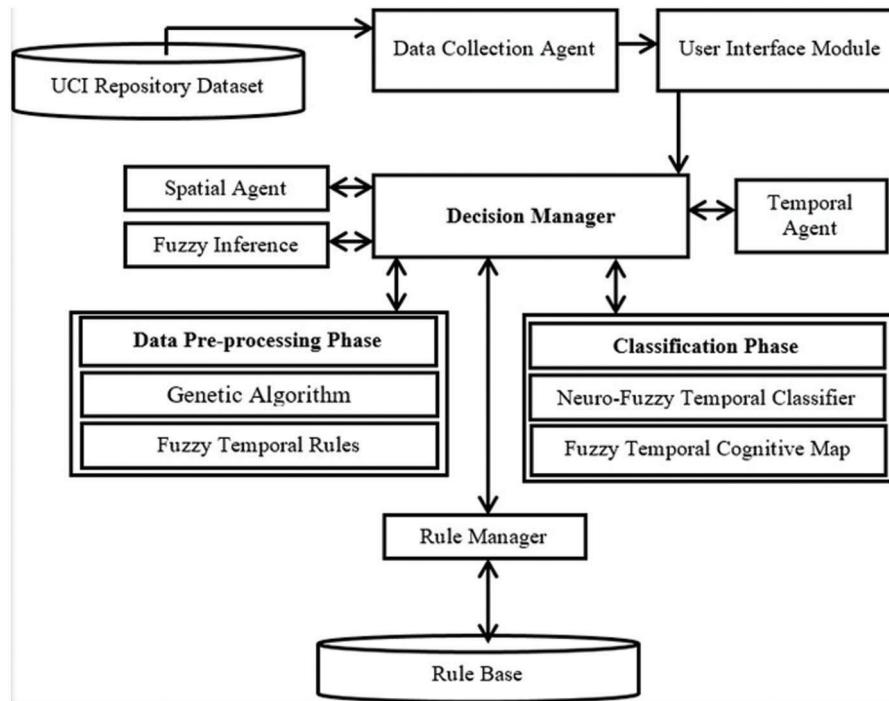
Figure 1: Overall system architecture

The Classification Phase consists of two sub components such as neuro-fuzy temporal classifier and fuzzy cognitive map that are useful for classifying the data effectively. The Rule manager is responsible for supplying the suitable and necessary rules to the decision manager in crucial time to make a decision. The Rule base contains the various rules that are useful for diabetes classification.

4 Proposed Work

In this paper, a new medical expert system has been developed for predicting the diabetic disease and the severity by introducing a new Fuzzy-Temporal Rule and Weighted Average Genetic Algorithm based Feature Selection Algorithm (FTWGA-FSA) for effective feature selection and the Neuro-Fuzzy Temporal Rule based Classification (NFTRCA) algorithm and the existing Fuzzy Cognitive Map (FCM) for effective classification. This section explains in detail about the proposed medical expert system.

4.1 Feature Selection

In this thesis, feature selection was carried out based on the UCI Diabetic Repository. Later, new and additional features were added by consulting with domain experts to suit Indian conditions. This work proposes a new genetic aware feature selection method to select useful features. The major benefit of this method is to identify the useful features for predicting Type-1 and Type-2 diabetes. Moreover, the user's queries are to be accepted along with a filled questionnaire for completing their request. According to their registered information, the feature selection method categorizes the relevant features of Type-1 and Type-2 diabetes.

Normally, people with diabetes have either TYPE1 diabetes, caused by a lack of insulin or TYPE2 diabetes, in which case they have very little insulin. TYPE1 diabetics have problems with insulin production and their cells are, as a result, unable to absorb sugar (glucose). However, sugar is an essential component for every human being since it is necessary to produce energy. TYPE2 diabetes can develop at any age, chiefly for adults. In this case, the damage is great and must be attended to on time. The common symptoms of the diabetes are including fatigue, hunger, dry mouth, blurred vision and itchy skin. Especially, urination, losing weight and blurry vision are identified as important Type-1 diabetes symptoms. Yeast infection is identified as important additional symptom for Type-2 diabetic. These symptoms are very common for men and women. The newly developed feature selection method works by using the below steps.

Algorithm:

Fuzzy-Temporal Rule and Weighted Average Genetic Algorithm based Feature Selection

Input: Patient details with all features, mutation, crossover, population and time.

Output: Optimal features F_i for the specific time interval.

Step 1: Let assume that the population are positioned as randomly with the size of each chromosome as 64.

Step 2: Symptoms represents by using 48 bits

Step 3: Time constraints represent by using 16 bits   // Bit values are either '0' or '1'

Step 4: Weights are initialized as follows:

   $Wt_1 = 0.65$, $Wt_2 = 0.45$, $Wt_3 = 0.75$, $Wt_4 = 0.35$

Step 5: Read all 'N' records from dataset that are available as training dataset.

Step 6: Read the crossover probability value 'CP_c'.

Step 7: Read the mutation probability value 'MP_m'.

Step 8: Read the starting time 'st' and ending time 'et' for each activity as time interval.

Step 9: For each chromosome in the current population valid during the interval [st, et]

   i. Use an equal crossover value with the probability CP_c.
   ii. Use a mutation operator to the chromosome with the mutation probability P_m.
   iii. Set time limits as interval [st, et]
   iv. Apply the following formula for calculating the fitness value (FIT).

   FIT = [Wt_1 * Count of ones in first 48 bits + Wt_2 * Count of ones in first 48 bits +

   Wt_3 * Count of ones in first 16 bits + Wt_4 * Count of ones in first 16 bits]/(W1+W2+W3+W4)

Step 10: If (CF–PF < 0.0001) then // CF-Current Fitness, PF-Previous Fitness

   Apply Fuzzy-Temporal Rules and select the features to fit into the new population for forming new parents.

   Else

   Feature will not be selected.

Step 11: If there is more population then

   Go to step 9.

   Else

   Return (FS_i, st, et)

Step 12: End

The proposed feature selection algorithm called Fuzzy-Temporal Rule and Weighted Average Genetic Algorithm based Feature Selection Algorithm (FTWGA-FSA) is used for identifying the suitable and the most contributed features that are useful for making effective decision over the testing process. The most contributed features have been identified and selected by using the proposed fuzzy temporal rules, fitness functions, weights assigned for the different attributes. Here, the features are considered as chromosomes and also perform the basic genetic operations such as population generation, mutation and crossover processes. Moreover, the feature values are converted into 48 bits and the time is converted into 16 bits for performing the mutation and crossover operations along with the randomly fixed weights. Here, the weights are playing major role to make decision on medical records. Then, it calculates the fitness value according to the feature weightages. The last decision on features is taken by using the fitness value and fuzzy temporal rules. Finally, the most contributed features are to be selected and returned as output. The selected features are considered as input for the disease prediction process of the proposed Neuro-Fuzzy Temporal Classifier. The proposed feature selection algorithm is different from the available feature selection algorithms as it applies the fuzzy logic, temporal constraints, genetic operations and weight consideration.

4.2 Neuro-Fuzzy Temporal Classification

This section provides a detailed study about the newly developed classifier called Neutral Network Classifier with Fuzzy logic and Temporal constraints (NFTC) which is proposed based on the neural network and fuzzy temporal rules. The major benefit of this work is that it employs time limitation for feature selection, classification and temporal reasoning based decision making. The proposed classifier works with the following steps.

Input: Set of Diabetic records D_i, Threshold δ.

Parameters used: Limits for Trapezoidal Membership function [L_i, U_i]

Output: Set of fuzzy temporal rules [$FTR(i)$, $i = 1 \ldots m$, $t = [TS_j, TE_i, j = 1 \ldots n]$]

$\qquad\qquad\qquad\qquad$ // m is the number of rules

Phase 1: Training

Step 1: Read the first record from the dataset DS

Step 2: Read the first feature set *(FS)* // Feature Set FS

Step 3: While (DS_i != *End of File*)

Begin

Step 4: Read the next record from the dataset DS_i

Step 5: Call Artificial Neural Network Tree to train the feature set *FS* for the specific records.

Step 6: Assign the weightage for the specific record set with time constraints *(st, et)*

Step 7: If the patient (record) has symptom of common diabetic and age is less than 18 then

$\qquad$ Update the result of the patient record is (RS_1) with diabetic symptom in low age

$\qquad$ else

$\qquad$ Update the result of the patient record is (RS_1) with diabetic symptom in high age

Step 8: If the record has symptom for TYPE-1 diabetic and age is less than 18 then

$\qquad$ Apply the fuzzification process with the inputs of record set, weight, time constraints and stored as result in R_3

(Continued)

(continued)

Else if the patient (record) has symptom for TYPE-1 diabetic and age is greater than or equal to 18 then

Apply the fuzzification process with the inputs of record set, weight, time constraints and stored as result in R_4

Else if the patient (record) has symptom for TYPE-2 diabetic and age is less than 18 then

Apply the fuzzification process with the inputs of record set, weight, time constraints and stored as result in R_5

Else if the patient (record) has symptom for TYPE-2 diabetic and age is greater than or equal to 18 then

Apply the fuzzification process with the inputs of record set, weight, time constraints and stored as result in R_6

Step 9: If the Old value of RS – New value of RS is less than the threshold δ then

Go to Step 10

else

Go to Step 3

Step 10: Return the New value of record (RS) with fuzzy temporal rules (FTR_i)

End

Phase 2: Testing

Step 1: Read the user query through patient disease status enquiry form.

Step 2: Retrieve the relevant feature set based on the user request with the time consideration.

Step 3: If the user is authenticated user or not for the application by verifying the password that has been entered by the patient (user) Then

Go to Step 4

Else

Display the registration page with login credentials as new user.

Step 4: Send the user/patient request to the decision manager for selecting the suitable fuzzy temporal rules from rule base with the help of rule manager for making effective decision.

Step 5: Apply fuzzy inference through decision manager for making final decision over the specific record with the help of expert advice.

Step 6: Apply the Fuzzy Cognitive Map (FCM) along with necessary inputs what we have received as output from the existing steps.

Step 7: Display the decision according to the FCM result as a final decision/result for the user query as report and form.

This algorithm makes decisions using fuzzy rules with help of the inference engine and expert advice. The fuzzy rules are used to perform a deductive inference are as follows:

A sample set of rules for the diabetic dataset is presented below:

Rule 1:

IF the patient's blood sugar value is below 120 after a meal AND HbA1c< 6.0% AND [et –st] > 90 days **THEN**

    IF

        (Polydipsia < Threshold AND

        Polyphagia < Threshold AND

        Polyuria < Threshold AND

        Blurred vision < Threshold AND

        Extreme tiredness < Threshold AND

        Genital Itching < Threshold AND

        Slow healing of wounds < Threshold AND

        Unexplained weight loss < Threshold)

    THEN

    Diabetics = False

    END IF

The patient does not require any treatment.

END IF

Rule 2:

IF the patient's insulin value >= 200 AND insulin value < 250 AND HbA1c >= 8.0% AND HbA1c < 9.0% AND [t2 – t1] >= 360 days **THEN**

    Diabetics = High Medium

// It requires 5 units of regular insulin dosage, pre-breakfast and pre-dinner.

END IF

Rule 3:

IF the patient's insulin value > 250 AND HbA1c >=10.0% AND [et – st] >= 360 days **THEN**

    IF

        Serious loss of appetite AND nausea and vomiting OR

        Serious loss of appetite AND high temperature OR

        Nausea and vomiting AND high temperature OR

        Pain in the stomach AND serious loss of appetite OR

(continued)

Pain in the stomach AND nausea and vomiting OR

Pain in the stomach AND high temperature OR

Fruity or chemical smell on the breath AND serious loss of appetite OR

Fruity or chemical smell on the breath AND nausea and vomiting OR

Fruity or chemical smell on the breath AND high temperature OR

Fruity or chemical smell on the breath AND pain in the stomach

THEN

END IF

Diabetics = Very High

END IF

// It requires 8–9 units of regular insulin dosage pre-breakfast and pre-dinner.

Rule 4:

IF Blood_glucose after 2 h of eating > 7.8 mmol/l AND [t2 – t1] >= 5 Years AND Parents and Ancestors with diabetic> 1 **THEN**

Type1 Diabetic = True

// predicted, based on history of family members.

END IF

Rule 5:

IF Starting stage of Diabetes = True AND Skin disorder AND Sudden weight loss OR Loss of muscle mass AND Giddiness AND [t2 – t1]>= 365 days

THEN

Type2 diabetic = Low

// Predicted based on symptoms

END IF

Fuzzy rules are developed using expert advice and medical parameters. The full set of rules designed in this thesis is used by the inference engine to make effective decisions by applying forward chaining inference. Further, it builds a discriminate network to perform rule matching and rule firing. The necessary rules are generated by considering the medical experts' advices and the fuzzy intervals and temporal constraints for making final decision on medical dataset that are extracted from the standard benchmark dataset and streaming dataset. All these fuzzy temporal rules are applied in the proposed neural temporal fuzzy classification for making effective decisions on record sets.

5 Results and Discussion

In this work, experiments were conducted using both benchmark datasets and real datasets collected from hospitals. In addition, domain experts were consulted to frame rules. The benchmark dataset used in this work was obtained from the University of California, Irvine (UCI) Repository. It consists of 2 classes for indicating the diabetes disease with Type-1 and Type-2 categories with 768 records. These are all the records of the patients who crossed the age 21 years and collected from UCI repository dataset along with records such as pregnant count, two hours plasma glucose concentration, blood pressure, skin thickness, serum insulin for two hours, body mass index value and age. Among them, five features such as glucose, blood pressure, skin thickness, body mass index value and insulin values are considered and selected by the proposed feature selection method. The live records collected for evaluating the proposed model in this work is presented in Tab. 1.

Table 1: Details of patients

Parameter	Details
Total no. of patients	250
No. of males & females	110 & 140
Age	10–65 yrs
Duration of observation	3 months to 6 years
Duration of follow-up	Weekly once

The follow-up procedure is also to be done and it varies from patient to patient depending upon the medical history. In this work, the medical history of the patients is also considered that are maintained from 2008 to 2014. The patients were considered from the cities such as Bangalore, Chennai, Kolkata, New Delhi and Mumbai. The city wise computational time analysis is demonstrated in Fig. 2.

From Fig. 2, it is proved that the training and testing time vary significantly from city to city according to the number of records considered. The training time and the testing time of the cities like Mumbai and Kolkata are less than the cities including Chennai, Bangalore and New Delhi. The training and testing times are optimal due to the application of genetic algorithm with the consideration of time.
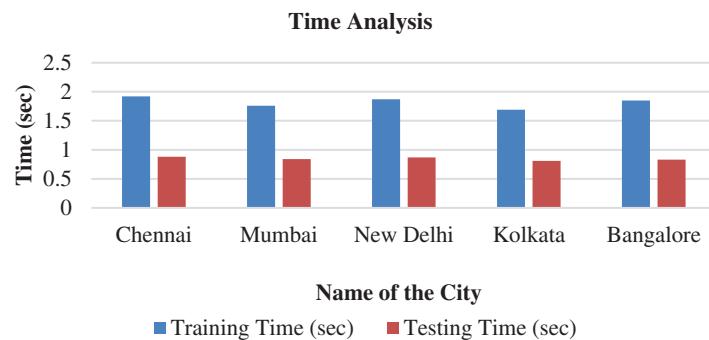
Figure 2: Time analysis

Fig. 3 demonstrates the performance of the proposed classifier with the consideration of full dataset and the consideration of selected features. Here, the proposed neural network classifier with the consideration of

fuzzy logic and time constraints are evaluated by conducting five different experiments with different number of records. The performance is calculated for classification accuracy on the input dataset.
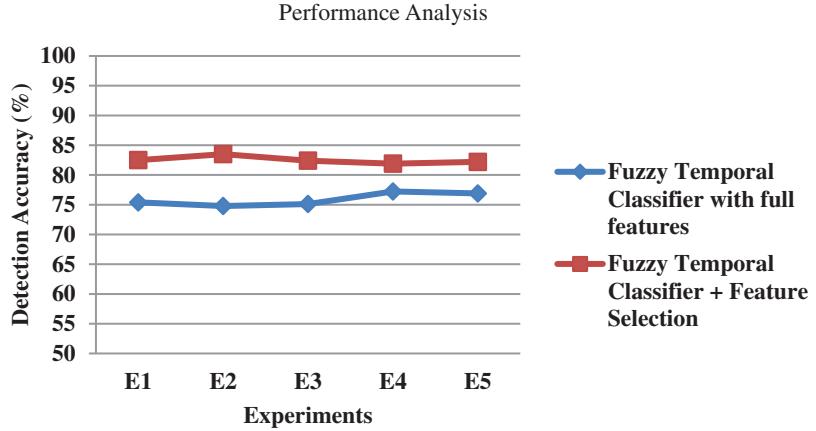
Figure 3: Performance analysis

Fig. 3 proves the effectiveness of the proposed neural classifiers with the consideration that the selected featured dataset is better than the proposed neural classifier with full features. This is because the contributory features with temporal constraints converge faster at the correct decision values. On the other hand, the presence of additional features confuses the decision-making process, leading to increased decision time and reduced accuracy.

Fig. 4 shows the performance comparative analysis between the fuzzy classifier with time and without time constraints. Here, five different experiments have been done with different sets of patient records collected from diverse hospitals in different Indian cities.
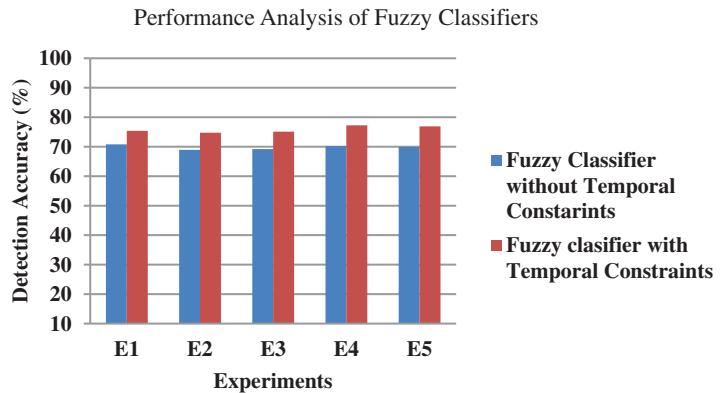
Figure 4: Performance comparative analysis of the fuzzy classifier

From Fig. 4, it can be observed that the proposed fuzzy classifier with temporal constraints is better than the one without temporal constraints, with respect to classification accuracy. This is owing to the fact that the use of temporal constraints helps analyze the past effectively to form rules to perform temporal reasoning which provides learning and prediction features. Moreover, the temporal constraint applied with the fuzzy classification algorithm is able to perform qualitative temporal reasoning which increases decision accuracy.

Fig. 5 shows a comparative analysis of fuzzy classifiers without and with time constraints, fuzzy classifier with feature selection and time and the fuzzy classifier without feature selection and time constraints. Five experiments were conducted with different sets of patient records from the standard benchmark dataset.
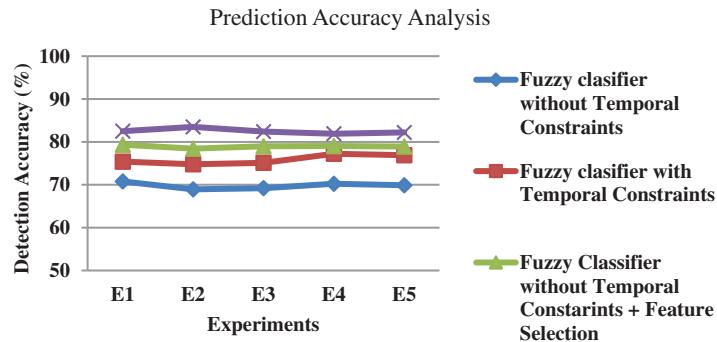
Figure 5: Prediction accuracy analysis

From Fig. 5, it can be observed that the proposed classifier with feature selection fared well than the proposed classifier without feature selection and time constraints. This is because the use of selected features with time constraints provides deductive inferences on learned rules leading to increased detection accuracy. Moreover, the results obtained from the proposed classifier have been tested with live dataset and validated using expert advice from domain experts. The experiments conducted in this work show that there is no difference between decisions made by this system and expert opinions. Hence, the proposed algorithm is far more accurate and capable of making recommendations for planning treatment.

Fig. 6 shows a comparative analysis between the proposed disease prediction system and the existing disease prediction systems that are developed by using the classifiers namely C4.5, Fuzzy Temporal Cognitive Map (FTCM) [21], Fuzzy Min-Max- Genetic Algorithm (FMM-GA) [2], and Temporal Fuzzy Min-Max – Particle Swarm Optimization (TFMM-PSO) [19], Wu et al. [21] and Ahmed et al. [23]. In this model, five experiments were conducted with different sets of records such as 50, 100, 150, 200 and 250 from the standard dataset and the live dataset. The set of records are collected from different hospitals located in assorted Indian cities.
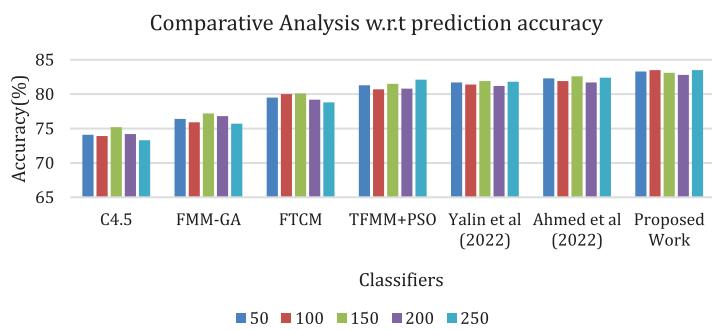
Figure 6: Comparative analysis w.r.t prediction accuracy

From Fig. 6, it is evident that the proposed prediction system is better than the existing disease prediction systems that are developed by using the classifiers such as C4.5, Fuzzy Min-Max- Genetic Algorithm (FMM-GA), Fuzzy Temporal Cognitive Map (FTCM) and Temporal Fuzzy Min-Max – Particle Swarm

Optimization (TFMM-PSO), Wu et al. [21] and Ahmed et al. [23]. In this, the performance of the proposed model has achieved better performance in all the five experiments that are considered for different patient records. Random records were selected for conducting experiments and also for evaluating the proposed model. All the existing works applied fuzzy logic, neural network, temporal constraints and optimization techniques for performing prediction process. Even though, they failed to attain the current requirements. The proposed model applied effective fuzzy temporal rules for predicting the diseases and the severity level accurately. The reason for the improvement is the incorporation of genetic algorithm, temporal logic, fuzzy logic, neural network and feature selection process. The proposed disease prediction system performs well than the existing systems in terms of disease prediction accuracy.

6 Conclusion and Future Work

The new disease prediction system has been developed and implemented for diagnosing diabetes effectively by applying neural network with fuzzy logic and time constraints, genetic algorithm-based feature selection method. This newly proposed classifier incorporates the fuzzy logic and temporal constraints that are helpful for making effective decision on patient records. The newly developed genetic algorithm-based feature selection algorithm selects the most contributed features that are helpful for enhancing the prediction accuracy. The proposed disease prediction system achieved better disease prediction accuracy than the existing systems that are available in the literature. This work can be enhanced further with the introduction of multi-neural network along with the consideration of fuzzy logic and temporal constraints.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] Y. Li, J. L. Wang, Z. H. Tian, T. B. Lu and C. Young, "Building lightweight intrusion detection system using wrapper-based feature selection mechanisms," *Computers & Security*, vol. 28, no. 6, pp. 466–475, 2009.

[2] S. Ganapathy, R. Sethukkarasi, P. Yogesh, P. Vijayakumar and A. Kannan, "An intelligent temporal pattern classification system using fuzzy temporal rules and particle swarm optimization," *Sadhana*, vol. 39, no. 2, pp. 283–302, 2014.

[3] R. K. Nowicki and J. T. Starczewski, "A new method for classification of imprecise data using fuzzy rough fuzzification," *Information Sciences*, vol. 414, pp. 33–52, 2017.

[4] M. Pota, M. Esposito and G. D. Pietro, "Designing rule-based fuzzy systems for classification in medicine," *Knowledge-Based Systems*, vol. 124, pp. 105–132, 2017.

[5] M. F. Mohammed and C. P. Lim, "Improving the fuzzy min-max neural network with a k-nearest hyperbox expansion rule for pattern classification," *Applied Soft Computing*, vol. 52, pp. 135–145, 2017.

[6] H. R. Singh, S. K. Biswas and B. Purkayastha, "A Neuro-fuzzy classification technique using dynamic clustering and GSS rule generation," *Journal of Computational and Applied Mathematics*, vol. 309, pp. 683–694, 2017.

[7] I. Kavakiotis, O. Tsave, A. Salifoglou, N. Maglaveras, I. Vlahavas *et al.,* "Machine learning and data mining methods in diabetes research," *Computational and Structural Biotechnology Journal*, vol. 15, pp. 104–116, 2017.

[8] B. Sasikumar, D. Naveenraju, K. Anand, S. Hariharan, P. Sudhakaran *et al.,* "Diabetes prediction using sensors by analysing skin temperature," *Journal of Engineering Science and Technology*, vol. 15, no. 2, pp. 1357–1370, 2020.

[9] M. Kusy, "Fuzzy c-means-based architecture reduction of a probabilistic neural network," *Neural Networks*, vol. 108, pp. 20–32, 2018.

[10] E. Cordelli, G. Maulucci, M. D. Spirito, A. Rizzi, D. Pitocco *et al.,* "A decision support system for type 1 diabetes mellitus diagnostics based on dual channel analysis of red blood cell membrane fluidity," *Computer Methods and Programs in Biomedicine*, vol. 162, pp. 263–271, 2018.

[11] R. Cheruku, D. R. Edla, V. Kuppili and R. Dharavath, "RST-BatMiner: A fuzzy rule miner integrating rough set feature selection and bat optimization for detection of diabetes disease," *Applied Soft Computing*, vol. 67, pp. 764–780, 2018.

[12] A. Umamageswari, N. Bharathiraja and D. S. Irene, "A novel fuzzy c-means based chameleon swarm algorithm for segmentation and progressive neural architecture search for plant disease classification," *ICT Express*, 2021. ISSN 2405-9595, https://doi.org/10.1016/j.icte.2021.08.019.

[13] S. Cui, D. Wang, Y. Wang, P. W. Yu and Y. Jin, "An improved support vector machine-based diabetic readmission prediction," *Computer Methods and Programs in Biomedicine*, vol. 166, pp. 123–135, 2018.

[14] F. d. A. T. de Carvalho, E. C. Simões, L. V. C. Santana and M. R. P. Ferreira, "Gaussian kernel c-means hard clustering algorithms with automated computation of the width hyper-parameters," *Pattern Recognition*, vol. 79, pp. 370–386, 2018.

[15] P. M. Murphy and D. W. Aha, "UCI repository of machine learning databases," in *Irvine*, CA: University of California, Department of Information and Computer Science, 1995.

[16] U. Kanimozhi, S. Ganapathy, D. Manjula and A. Kannan, "An intelligent risk prediction system for breast cancer using fuzzy temporal rules," *National Academy Science Letters*, vol. 42, no. 3, pp. 227–232, 2018.

[17] M. S. Nawaz, B. Shoaib and M. A. Ashraf, "Intelligent cardiovascular disease prediction empowered with gradient descent optimization," *Heliyon*, vol. 7, no. 5, pp. e06948, 2021.

[18] M. M. Ali, B. K. Paul, K. Ahmed, F. M. Bui, J. M. W. Quinn *et al.,* "Heart disease prediction using supervised machine learning algorithms: Performance analysis and comparison," *Computers in Biology and Medicine*, vol. 136, pp. 104672, 2021.

[19] A. Quteishat, C. P. Lim and K. S. Tan, "A modified fuzzy min–max neural network with a genetic-algorithm-based rule extractor for pattern classification," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 40, no. 3, pp. 641–650, 2010.

[20] R. Sethukkarasi, S. Ganapathy, P. Yogesh and A. Kannan, "An intelligent neuro fuzzy temporal knowledge representation model for mining temporal patterns," *Journal of Intelligent & Fuzzy Systems*, vol. 26, no. 3, pp. 1167–1178, 2014.

[21] Y. Wu, Q. Zhang, Y. Hu, K. Sun-Woo, X. Zhang *et al.,* "Novel binary logistic regression model based on feature transformation of XGBoost for type 2 diabetes mellitus prediction in healthcare systems," *Future Generation Computer Systems*, vol. 129, pp. 1–12, 2022.

[22] D. Acharya and D. K. Das, "An efficient nonlinear explicit model predictive control to regulate blood glucose in type-1 diabetic patient under parametric uncertainties," *Biomedical Signal Processing and Control*, vol. 71, no. A, pp. 103166, 2022.

[23] U. Ahmed, G. F. Issa, M. A. Khan, S. Aftab, M. F. Khan *et al.,* "Prediction of diabetes empowered with fused machine learning," *IEEE Access*, vol. 10, pp. 8529–8538, 2022.