

Self Organising Fuzzy Logic Classifier for Predicting Type-2 Diabetes Mellitus using ACO-ANN

Ratna Patil¹

Computer Engineering Department
Noida Institute of Engineering and
Technology, Greater Noida
Uttar Pradesh, India

Dr. Sharvari Tamane²

Professor, Information Technology
Jawaharlal Nehru College of
Engineering, Aurangabad
Maharashtra, India

Dr. Kanishk Patil³

Resident Doctor
J J Hospital, Mumbai
Maharashtra
India

Abstract—In today's digital world, a dataset with large number of attributes has a curse of dimensionality where the computation time grows exponentially with the number of dimensions. To overcome the problem of computation time and space, appropriate method of feature selection can be developed using metaheuristic approaches. The aim of this work is to investigate the use of ant colony optimization with the help of neural network to select near optimal feature subset and integrate it with the self-organizing fuzzy logic classifier for improving the recognition rate. The proposed fuzzy classifier derives prototype from the collected data through an offline training process and uses it to develop a fuzzy inference system for classification. Once trained, it can continuously learn from streaming data and later adapts the changing facts by updating the system structure recursively. The developed model is not based on predefined parameters used in the data generation model but is derived from the empirically observed data.

Keywords—Ant colony optimization; feature selection; fuzzy logic classifier; self organizing; type-2 diabetes mellitus

I. INTRODUCTION

Machine learning algorithms are widely used in medical field. Several classification and clustering techniques are useful for disease prediction. Nowadays, we are living in the era of BIG-DATA, so classification algorithms have got importance in the research. Conventional classification algorithms are trained on static datasets. Once the classifiers are trained, no modification is possible in their configuration.

Most of the classifiers are developed when data is not available in the largescale. To overcome this issue, online approach can be used for developing classifiers. In this approach classifiers are constantly learning from new instances [1]. They also store key information and disposes already processed instances. Variation in the pattern in nonstationary situations is considered by evolving system structure and recursively modifying meta – data. Pattern may change when the data availability is continuously increasing, and conventional offline approaches will not consider this fact. However, it is not feasible to learn online from the beginning because setting the system with the presented static data in an offline manner can assure improved performance. SOF classifier is trained in two stages. The classifier is trained offline with the presented static data and fuzzy rule-based system is obtained in the first stage. During the second online training stage, the fuzzy rule-based system recognized over

the offline training process will be restructured after the new instances are processed to track the possible changes in the data pattern [1] [2].

Our study is based on the development of predictive model for Type-2 diabetes Mellitus (T2DM). Diabetes is a chronic illness accountable for increasing number of deaths every year around the world. It typically progresses after years of insulin resistance, or prediabetes. Prediabetes can progress to diabetes when liver and muscle cells become more and more insulin resistant (they don't respond properly to this internal insulin signal) and have growing problems of increased sugar in the bloodstream. Physicians suggest that if prediabetes is recognized early its progression can be halted. As per physician's advice, vegetables and fruits should form the bulk of the diet, losing weight and exercising are the best ways to stop prediabetes from becoming Type 2 diabetes. If prediabetic stage is not treated it can lead to type 2 diabetes and complications like heart disease, renal disease and stroke. Machine learning algorithms can be used to develop a predictive model for T2DM [3]. If prediabetic stage is detected on time, then change in lifestyle can prevent the diabetes.

I have selected this topic because the current approaches have often classified the patients as either having diabetes or a healthy individual, ignoring the prediabetic state of a patient. This leads to increased incidence of diabetes if no proper preventive measures are taken on time. It can be curbed at the prediabetic stage by introducing lifestyle changes.

Most of the work available in the literature has been done on PIMA dataset of the UCI repository. This dataset has 8 attributes and only two labels (Healthy / Diabetic). There is no mention of prediabetic stage and hence I have selected this topic.

Literature review shows that most of the developed traditional classifiers are offline [3] [4] [5] Once they are trained on dataset which is static in nature then no further modification is possible in the structure of classifier model. These traditional methodologies require users to predefine various kinds of parameters to obtain promising result. In real cases this prior knowledge may not be available. The recent study shows that most of the work has been done on offline data processing and retraining is required when data pattern changes. In our study, we have overcome this limitation and

taken into consideration changing data pattern while developing the predictive model.

This paper is presented by writing five sections. Section one gives the introduction of reasons behind the selection of study. Section two discusses the empirical data analytics operators in brief. Section three discusses the feature selection method. The pseudocode of stage 1 (static offline training) and stage 2 classifier (self-evolving training) has also been discussed in section three. Section four comprises of the analysis and experimental results. Section five contains conclusion of the experimental study.

II. NOTIONAL BASIS

Below mentioned statistical calculations are required for the proposed method [2]:

- 1) Cumulative proximity,
- 2) Unimodal density,
- 3) Multimodal density,

4) Along with these proximity and density calculation, we need to compute their recursive forms for streaming data processing.

a) Cumulative proximity- The cumulative proximity between two data points is computed by the equation 1.

$$\pi_k(x_i) = \sum_{j=1}^k d^2(x_i, x_j); i = 1, 2, \dots, k \quad (1)$$

Where, $d(x_i, x_j)$ denotes the distance between two points which can be measured as Euclidean, Makowski or Cosine.

b) Unimodal density- This specifies the main data pattern in our proposed method and computed by the equation 2.

$$D_k(x_i) = \frac{\sum_{l=1}^k \pi_k(x_l)}{2K\pi_k(x_i)} = \frac{\sum_{l=1}^k \sum_{j=1}^k d^2(x_l, x_j)}{2k \sum_{j=1}^k d^2(x_i, x_j)}; \quad (2)$$

$i = 1, 2, \dots, k$

c) Multimodal density- This is calculated by the equation 3 where, f_i specifies the corresponding frequency of occurrence.

$$D_k^{MM}(u_i) = f_i D_k(u_i) = f_i \frac{\sum_{i=1}^k \pi_k(x_i)}{2K\pi_k(u_i)}; i = 1, 2, \dots, U_k \quad (3)$$

d) Recursive computation of densities - The recursive computation has a substantial role in the second stage of developing SOF classifier. We get well-designed recursive calculation forms, using the equation 4 where quantities can be updated by an effective means by storing only the key meta-parameters.

$$\pi_k(x_i) = K((x_i - \mu_k) \sum_k^{-1} (x_i - \mu_k)^T + X_k - \mu_k \sum_k^{-1} \mu_k^T) \quad (4)$$

Where, recursive definitions of global mean μ_k , covariance matrix \sum_k and X_k are computed by equation 5, 6 and 7, respectively.

$$\mu_k = \frac{K-1}{K} \mu_{K-1} + \frac{1}{K} x_K; \mu_1 = x_1 \quad (5)$$

$$X_K = \frac{K-1}{K} X_{K-1} + \frac{1}{K} x_K^T x_K; X_1 = x_1^T x_1 \quad (6)$$

$$\sum_K = \frac{K}{K-1} (X_K - \mu_K^T \mu_K) \quad (7)$$

III. PROPOSED METHODOLOGY

In this experimental study we have investigated the use of ant colony optimization with the help of artificial neural network (ACO-ANN) to select near optimal feature subset and integrated it with the self-organizing fuzzy logic classifier for improving the recognition rate.

In this study, we have proposed Ant Colony Optimization method for feature selection [4] and ANN is used for implementing fitness function. The obtained feature subset is then further used for developing SOF classifier. The architecture diagram of proposed system is depicted in Fig. 1. Literature shows that there are various machine learning algorithms [6], metaheuristic algorithms [7] [8] and fuzzy logic methods [9] which can be used for feature selection. Based on literature survey feature selection using nature inspired algorithms Ant colony optimization gives higher accuracy [5]. ACO is inspired by the behaviour of ant colonies. When ants go for searching food, they deposit pheromone on the path. This odorous substance is used as a communication medium. The quantity of the placed pheromone depends on the distance of the food source. Other ants moving at random detects a laid pheromone and most likely the ant follows the same path. These ants will also place the pheromone. Consequently, the path which is used by more ants will be followed. Probability of path which ant decides to choose increases with the number of ants previously followed that path. An artificial ant builds solutions to an optimization problem. Hybrid algorithm was developed to select prominent features using ACO. An artificial ant can be used for selecting subset of features. Ants traverses the node which represents features to construct a subset. All the ants attempt to construct a subset. The conventional probabilistic transition rule for selecting features is used. The node is chosen based on the probability which is computed by equation 8.

$$p_{ij}^k = \begin{cases} \frac{\tau_{ij}^\alpha \eta_{ij}^\beta}{\sum_{c_{il} \in f(s^p)} \tau_{il}^\alpha \eta_{il}^\beta} & \text{if } c_{il} \in f(s^p) \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

Where, $f(s^p)$ represents the set of possible features which can be further added to the solution being constructed. τ_i denotes pheromone values and η_i denotes heuristic values related to feature $i = 1, 2, 3, \dots, n$. α and β are two constants which regulates the pheromone value and heuristic information respectively. At each iteration pheromone values τ_{ij} , is updated by all the ants and the updated values at every path (i, j) is computed by equation 9.

$$\tau_{ij} = (1 - \rho) \tau_{ij} + \sum_{k=1}^m \Delta \tau_{ij}^k \quad (9)$$

Where, ρ denotes rate of evaporation, m denotes number of an artificial ant and $\Delta \tau_{ij}^k$ is pheromone laid on path (i, j) by k^{th} artificial ant. Equation 10 computes $\Delta \tau_{ij}^k$.

$$\Delta \tau_{ij}^k = \begin{cases} \frac{Q}{f} & \text{if an ant used path } (i, j) \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

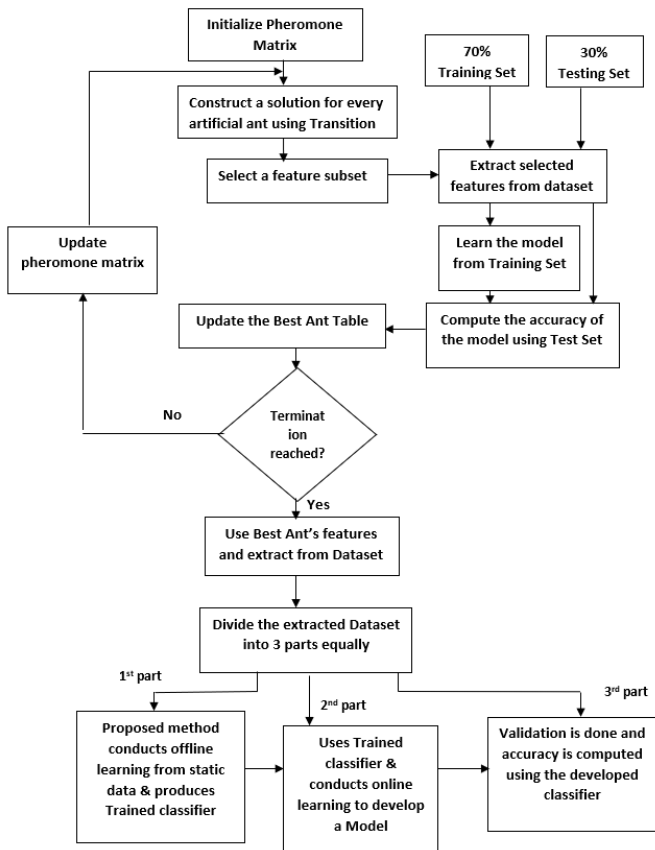


Fig. 1. Architecture of Proposed System.

Where, Q is constant, and f refers to the length of the tour by k^{th} ant. Integrating ACO with Artificial neural network (ANN) so that important features are extracted and then SOF is employed for classification. ANN has been employed for developing fitness function of ACO [8].

SOF is implemented in two stages. SOF will learn from the static dataset available at the beginning and recognize prototypes from each class independently to form 0-order fuzzy rule created on the acknowledged samples per class [1]. There is no impact of training processes of one class on another. The training process is led on data examples of the c^{th} class ($C = 1, 2, 3, \dots, N$). Prototypes are recognized based on the distributions of the data samples and its density. Firstly, multimodal densities are computed using all distinct instances in a class by the equation 4.

Pseudocode for First stage of classifier (static offline)

1. Compute D^{MM} over all the unique samples for each class by equation 3.
2. Find r_1 by equation 11 $\{D_{K^c}^{\text{MM}} u_i^c\}$.
3. Delete r_1 from $\{u\}_{U_K^c}$.
4. $K := 1, \{r\} := r_1, D_{K^c}^{\text{MM}}(r) = D_{K^c}^{\text{MM}}(r_1)$
5. Repeat following until $u_k^c - k$ is not empty
 $K++;$

Find r_k and delete it from $\{u\}_{U_K^c}$
Add r_k to r i.e. $\{r\} = \{r\} + r_k$
 $\{D_{K^c}^{\text{MM}}(r)\} \leftarrow \{D_{K^c}^{\text{MM}}(r)\} + D_{K^c}^{\text{MM}}(r_k)$

6. Recognize $\{p\}_0$ by equation 12 and form data clouds around it.
7. Compute the center of the data clouds denoted by $\{\varphi\}_0$ and compute D^{MM} around it.
8. Find $\{\varphi\}^{\text{neighbours}}$ and $\{p\}^c$ by equation 13 and 14 respectively.
9. Generate c^{th} fuzzy rule with $\{p\}^c$

$$r_1 = \arg \max_{i=1,2,\dots,U_K^c} (D_{K^c}^{\text{MM}}(u_i^c)) \quad (11)$$

$$\text{if } \left\{ (D_{K^c}^{\text{MM}}(r_i) > D_{K^c}^{\text{MM}}(r_{i+1})) \text{ and } (D_{K^c}^{\text{MM}}(r_i) > D_{K^c}^{\text{MM}}(r_{i-1})) \right\} \text{ then } r_i \in \{p\}_0 \quad (12)$$

$$\text{if } (d^2(\varphi_i, \varphi_j) \leq G_{K^c}^{c,L}) \text{ then } (\varphi_j \in \{\varphi\}_i^{\text{neighbours}}) \quad (13)$$

$$\text{if } (D_{K^c}^{\text{MM}}(\varphi_i) > \max_{\varphi \in \{\varphi\}_i^{\text{neighbours}}} (D_{K^c}^{\text{MM}}(\varphi))) \text{ then } (\varphi_i \in \{p\}^c) \quad (14)$$

The data examples are ordered by their mutual distances and values of multimodal density and stored in $\{r\}$. Let r_1 be the highest multimodal density. r_2 is recognized as the instance with the minimum distance from r_1 and r_3 is identified based on the minimum distance from r_2 . Reiterate in this way and build the list $\{r\}$. Multimodal densities are ordered and specified by $\{D_{K^c}^{\text{MM}}(r)\}$. Prototypes denoted by $\{P\}_0$ are then recognized as the local maxima of the ordered multimodal densities. Centers of the data cloud is computed and denoted by $\{\varphi\}_0$. $\{\varphi\}_i^{\text{neighbours}}$ denotes the collection of the centers of the adjacent data clouds. After identifying all the representative prototypes of the c^{th} class fuzzy rules are constructed.

$$\text{if } (D_{K^c+1}(x_{K^c+1}^c) > \max_{p \in \{p\}^c} (D_{K^c+1}(P))) \text{ or } (D_{K^c+1}(x_{K^c+1}^c) > \min_{p \in \{p\}^c} (D_{K^c+1}(P))) \text{ then } (x_{K^c+1}^c \in \{p\}^c) \quad (15)$$

$$\text{if } (\min_{p \in \{p\}^c} d^2(x_{K^c+1}^c, p)) > G_{K^c+1}^{c,L} \text{ then } (x_{K^c+1}^c \in \{p\}^c) \quad (16)$$

During the second phase of training, the classifier continues to update its configuration on a sample basis when new data is provided. Like the offline training phase the set of fuzzy rules of different classes are adapted. When the $K+1^{\text{th}}$ sample of c^{th} class is provided then the meta parameters $\mu_{K^c+1}^c, X_{K^c+1}^c$ and $\sum_{K^c+1}^c$ are computed by equations 5, 6 and 7, respectively.

Pseudocode for Stage-2 self-evolving training (online learning)

1. Repeat steps 2 – 7 until the new instance of C^{th} class $x_{K^c+1}^c$ is presented.
2. Compute $\mu_{K^c+1}^c, X_{K^c+1}^c$ and $\sum_{K^c+1}^c$ by equations 5, 6 and 7, respectively.
3. Compute D at $x_{K^c+1}^c$ and $\{p\}^c$;
4. Evaluate equations 15 and 16 and use Boolean OR between them.
5. If the result of OR operation between equations 15 and 16 is true then,
 $++ N^c; P_N^c = x_{K^c+1+1}^c; S_N^c = 1;$

$\{p\}^c = \{p\}^c + P_{N^c}^c ;$
 Else
 Compute $P_{n^*}^c ;$
 $P_{n^*}^c = \frac{S_{N^*}^c}{S_{n^*}^c + 1} P_{n^*}^c + \frac{1}{S_{n^*}^c + 1} x_{K^c+1}^c ;$
 $S_{n^*}^c = S_{n^*}^c + 1 ;$
 Endif
 6. ++ $K^c ;$
 7. Update the fuzzy rule.

IV. ANALYSIS AND EXPERIMENTAL RESULTS

For this study the dataset has been collected from the local hospitals. The dataset is having 32 features and 1071 instances. Class variable has three values as Diabetic, Prediabetic and Healthy patients. All the features are not useful for diabetes classification. ACO is used for selecting important features and objective function is developed using ANN for selecting near optimal feature subset. The selected features from the above ACO-ANN hybrid model are used for developing the Self organizing classifier. The proposed method uses Euclidean distance and cosine similarity for the implementation of classifier [10]. The dataset has been divided into three parts equally. Classifier conducts offline learning on the first part of static data. Then the proposed classifier conducts online learning from streaming data from second part of the collected dataset. classifier Third part is reserved for validation and testing purpose on unseen data. The developed classifier performs validation on testing data. ACO has selected Family History, Eating Fruits/Vegetables, PPG, FPG, Age, Feeling Hungry, Exercise, Frequent Urination, itchy skin and gender as important features in the diabetes detection. In this work values for ACO algorithm parameters are as per the Table I.

Confusion matrix on validation data for Euclidean distance and cosine similarity is shown in Fig. 2 and Fig. 3, respectively. The accuracy obtained is 86.27% in diabetes detection when distance type is Euclidean distance and 80.67% in the case of cosine similarity.

We have obtained the overall values for estimating the performance of the proposed classifier. Overall values of parameters obtained for the developed model are given in Table II. In our experimental study we have computed Accuracy, Error, Sensitivity, Specificity, Precision, False positive rate, F1 score, Matthew Correlation Coefficient and Kappa values to measure the performance. Users of the developed classifier would require the results of prediction. Practically, only the probability of correct predictions is not sufficient. Hence the set of statistical measures are computed to describe the developed classifier performances in various aspects. The following four terms are necessary to measure the various performance metrics:

- True positive (TP) - The data rows belong to the positive class and has been correctly predicted.
- False positive (FP) - The data rows belong to the negative class and has been incorrectly predicted as positive

- True negative (TN) - The data rows belong to the negative class and has been correctly predicted.
- False negative (FN) - The data rows belong to the positive class and has been incorrectly predicted as negative. The below mentioned measures qualify the performance of the prediction model and are used in medicine to measure the accuracy of analytical procedure.

TABLE I. ACO PARAMETERS AND VALUES

Parameter	Value
Initial Pheromone τ_0	0.1
Pheromone Exponential Weight α	1
Heuristic Exponential Weight β	1
Evaporation rate ρ	0.06
Constant Q	1

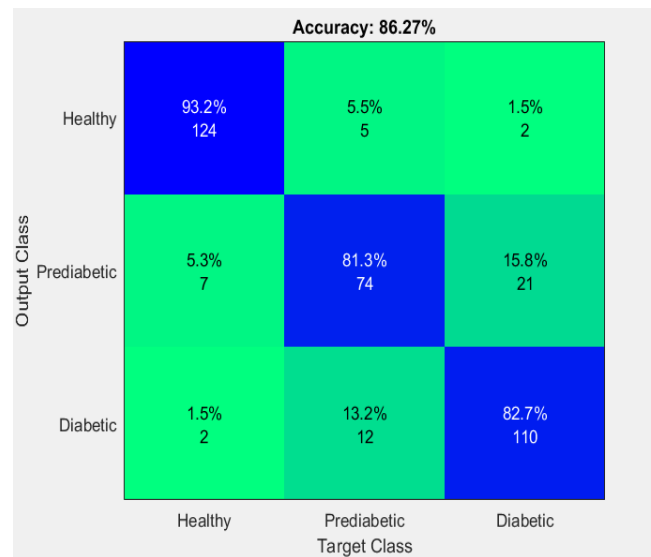


Fig. 2. Confusion Matrix on Validation Data (Euclidean Distance).

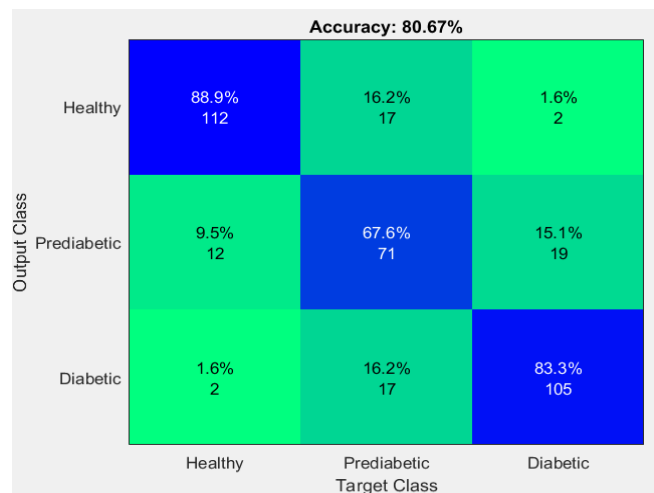


Fig. 3. Confusion Matrix on Validation Data (Cosine Similarity).

1) Accuracy: It is one of the most common evaluation measures to assess the performance of the model. Classification accuracy is given by the relation,

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} * 100$$

2) Error rate- Error rate gives incorrect predictions on total.

$$Error\ rate = (1 - accuracy) = \frac{(FP + FN)}{(TP + TN + FP + FN)}$$

3) Sensitivity: The sensitivity is used to certify the test capability of the model. it is the probability that the test is positive on malignant subjects. Sensitivity is obtained as the number of true positive results over the number of actual positive cases. It is referred as true positive rate. It is also called as recall.

$$Sensitivity = \frac{TP}{TP + FN} * 100$$

4) Specificity: It refers to the true negative rate. A good predictive model is one that maximizes both sensitivity and specificity. We want model to be both sensitive and specific.

$$Specificity = \frac{TN}{FP + TN} * 100$$

5) Precision: It is also called as positive predicted value. It is the proportion of positive samples that are truly positive. Higher precision is better. It efficiently computes an overall posterior probability and hence it is a meaningful performance measure.

$$Precision = \frac{TP}{TP + FP}$$

6) False positive rate (FPR): FPR is the ratio of false positives to the total number of negatives.

$$(1 - specificity) = False\ positive\ rate = \frac{FP}{FP + TN}$$

7) F1_score: It combines precision and recall using harmonic mean. Harmonic mean is used precision and recall both are expressed as proportions between zero and one. The worst F1 score value is zero and best is one.

$$F1_{score} = \frac{2 * recall * precision}{recall + precision}$$

8) Matthew Correlation Coefficient (MCC): MCC characterizes a confusion matrix with a single value. An MCC coefficient value 1 signifies a perfect prediction, 0 signifies no better than random prediction and -1 indicates total disagreement between prediction and observation.

$$MCC = \frac{(TP * TN - FP * FN)}{(TP + FP) * (TP + FN) * (TN + FP) * (TN + FN)}$$

9) Kappa coefficient: Kappa Statistic compares the accuracy of the system to the accuracy of a random system.

The Cohen's kappa coefficient is a statistical measure of inter-rater reliability. This coefficient is used to evaluate the classification accuracy. But it is normalized at the baseline of random chance on your dataset. In multi-class and imbalanced class problems, Cohen's kappa coefficient provides a good measure. Any kappa value less than 0.60 specifies inadequate agreement among the raters and less confidence should be shown in the study results.

$$Probability(Expected) = \frac{(TN + FP) * (TN + FN) + (FN + TP) * (FP + TP)}{Total * Total}$$

$$Kappa\ coefficient = \frac{Accuracy - Probability(Expected)}{1 - Probability(Expected)}$$

Multi-Class Confusion Matrix Output for the developed predictive model having distance type measure as Euclidean distance and cosine similarity is given in Table III.

TABLE II. STATISTICAL MEASURES OF DEVELOPED CLASSIFIER

S. No.	Parameters	Values (Euclidean Distance)	Values (Cosine similarity Measures)
1	Accuracy	86.27%	80.67%
2	Error rate	0.1373	0.1933
3	Sensitivity	85.31%	79.33%
4	Specificity	93.16%	90.49%
5	Precision	85.75%	79.95%
6	False Positive Rate	6.84%	9.51%
7	F1_score	0.8541	0.7992
8	Matthew Correlation Coefficient	0.7877	0.7039
9	Kappa	0.6912	0.5651

TABLE III. PREDICTIVE MODEL OUTPUT

	Class Value	True Positive	False Positive	False Negative	True Negative
Euclidean Distance	Healthy	124	9	7	217
	Prediabetic	74	17	28	238
	Diabetic	110	23	14	210
Cosine similarity	Healthy	112	14	19	212
	Prediabetic	71	34	31	221
	Diabetic	105	21	19	212

V. CONCLUSION AND FUTURE WORK

In this work, we have used metaheuristic approach for feature selection. ACO has been implemented to decide the features which can predict diabetes with higher accuracy. The proposed SOF Classifier algorithm does not depend on prior assumptions about the data generation model. But the proposed classifier learns from the static offline data and develops a prototype of Trained classifier. In the next stage the classifier conducts online learning from streaming data. The approach provides various types of similarity and distance

measures. We have used Euclidean distance and cosine similarity as distance measure for evaluating the performance of the developed model. It has been observed that accuracy has been 86.27% when the distance mode is Euclidean distance. The observed accuracy was 80.67% in the case of cosine similarity. If the person has been classified as prediabetic by the developed predictive model, then consult your doctor for further treatment. Early intervention has been revealed to delay, and in some cases prevent, the progression from prediabetes to diabetes. The proposed method will help the individual from developing T2DM.

As future work, the developed model shall be considered for the extension of more metacharacters in the construction of stage1 classifier and for other classification problems. Further application of first order fuzzy rules in the SOF classifier will increase the degrees of freedom and therefore enhance the efficiency of the developed model. Different types of distance metrics like Manhattan, Minkowski, Hamming distance will be used for the implementation of classifier and comparison with the existing result.

REFERENCES

- [1] G. Xiaowei and P. P. Angelov, "Self-Organising Fuzzy Logic Classifier," *Information Sciences*, vol. 447, pp. 36-51, June 2018.
- [2] P. Angelov, X. Gu and D. Kangin, "Empirical data analytics," *Int. J. Intell. Syst*, vol. 32, no. 12, p. 1261–1284, 2017.
- [3] R. Patil and S. Tamane, "A Comparative Analysis on the Evaluation of Classification Algorithms in the Prediction of Diabetes," *International Journal of Electrical and Computer Engineering*, vol. 8, no. 5, pp. 3966-3975, October 2018.
- [4] K. Monirul, M. Shahjahan and K. Murase, "Ant Colony Optimization Toward Feature Selection," in *Intech open*, February 20th 2013, pp. 1-44.
- [5] M. . F. Ganji and . M. . S. Abadeh, "A fuzzy classification system based on Ant Colony Optimization for diabetes," *Expert Systems with Applications*, pp. 14650-14659, 2011.
- [6] H. B. Alwan and K. R. Ku-Mahamud, "Feature selection and model selection algorithm," *International Journal of Mathematics and Computers In Simulation*, vol. 7, no. 5, pp. 406-414, 2013.
- [7] R. N. PATIL and S. TAMANE C, "Upgrading the Performance of KNN and Naïve Bayes in Diabetes Detection with Genetic Algorithm for Feature Selection," *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, vol. 3, no. 1, pp. 2456-3307, 2018.
- [8] P. Ratna and T. S. C., "PSO-ANN-Based Computer-Aided Diagnosis and Classification of Diabetes," in *Smart Trends in Computing and Communications*, *Proceedings of SmartCom 2019.*, 2019.
- [9] S. Zainuddin, F. Nhita and U. N. Wisesty, "Classification of gene expressions of lung cancer and colon tumor using Adaptive-Network-Based Fuzzy Inference System (ANFIS) with Ant Colony Optimization (ACO) as the feature selection," in *The 2nd International Conference on Data and Information Science*, 2019.
- [10] G. Xiaowei, A. Plamen and Z. Zhijin, "A Distance-Type-Insensitive Clustering Approach," *Applied Soft Computing*, Jan 2019.