International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Early Risk Prediction of Diabetes Categorization Using Fuzzy K-Means Clustering Algorithm

R. Sangeethapriya[1], S. Gomathi[2], B. Dhiyanesh[3*], J. K. Kiruthika[4], P. Saraswathi[5], K. Divya[6]

Submitted: 19/01/2024 **Revised**: 28/02/2024 **Accepted**: 05/03/2024

Abstract: Diabetes mellitus (DM) is a metabolic disease that primarily results in high blood glucose levels. There are distinct clinical types, with Type 1 and Type 2 being the most common forms of diabetes. A significant increase has been observed in the number of young people suffering from type 1 diabetes over the past few years for this reason. Diabetes can become chronic with a long latency period in childhood and adolescence, as the symptoms in the early stages can be vague. This can make timely detection and treatment complex, possibly leading to delayed treatment. It is important to detect or prevent diabetes early. It can cause many complications, and the prediction of diabetes is not accurate for further analysis using previous methods. We introduced the new proposed method using learning (ML) approaches to overcome the issues. Based on the Fuzzy K-means Clustering and Support Vector Machine (FKMC-SVM) for deciding the classification model for diabetic prediction using a standard dataset, for an accurate result, Initially collected, the diabetic dataset is from the standard repository, and the second step is pre-processing to reduce the imbalanced data, normalizing the values from the dataset using Z-Score normalization, and then selecting the features based on the margin values using Threshold Recursive Feature Elimination (TRFE) to eliminate the values from the pre-processing dataset based on the maximum threshold values of the recursive features in the dataset. Then the fuzzy-based method is used to decide diabetes using fuzzy logic to create interpretable models and to diagnose diabetes early based on these classifiers, FKMC and SVM, and to design fuzzy rules. FKMC refers to a collection of data points where the points in one location share similarities or connections but differ from those in another cluster. Additionally, optimizing support vector machines with larger datasets may provide more accurate results and predict the likelihood of diabetes in both Type 1 and Type 2. This combined algorithm, F-KMC-SVM, compares the precision, accuracy, recall, and F1-score of different ML techniques used to classify diabetic patients.

Keywords: *Feature Selection, Classification, Clustering, Fuzzy Rules, Machine learning, diabetes.*

1. Introduction

DM is a disease in which the body cannot produce sufficient insulin or use it effectively, resulting in elevated glucose (blood sugar) levels. There were 415 million diabetics worldwide in 2015. By 2040, this number will exceed 642 million. Additionally, 179 million people worldwide have undiagnosed diabetes. In addition, the WHO estimates that diabetes will account for 4.6 million global fatalities in 2030, making it the sixth largest cause of death. Diabetes is becoming a more challenging issue in the medical industry as the number of sufferers grows. Recent medical advancements have made it easier to diagnose diabetes early, yet 50% of individuals are still unaware they have the condition. Finding them might take ten years. Delays in therapy might result in severe side effects such as renal failure, blindness, hypertension, nerve damage, and stroke. Since diabetes mellitus is now an incurable condition, early detection and effective treatment are critical to its survival.

In order to diagnose diabetes, researchers have employed a variety of strategies and procedures over the years. ML is one of these technologies. ML was proven to be an essential tool in many industries, including medicine, during the fourth industrial revolution. Diabetes, for example, can be controlled, and early detection can save lives. In order to predict diabetes, various characteristics related to diabetes are given to the predictors. The Pima Indian diabetes dataset employs several ML clustering and classification algorithms to predict diabetes. Implicit computer or machine training is accomplished through ML. Different ML approaches enable effective knowledge collection by creating various classification and clustering models from the gathered datasets [21-23].

ML technology cannot explain diagnosis and data inference processes. Humans should understand technology. Another

[1]*Assistant professor, Department of Information Technology, Sona college of Technology, Salem 636 005, Tamil Nadu,*
E mail: priyabecse49@gmail.com
[2]*Assistant Professor, Department of Computer Science and Engineering, Dr. N.G.P Institute of Technology, Coimbatore 641048, Tamil Nadu.*
Email: mail2mathi86@gmail.com
[3]*Associate Professor, Department of Computer Science and Engineering, Dr. N.G.P Institute of Technology, Coimbatore 641048, Tamil Nadu.*
*Email: dhiyanu87@gmail.com**
[4]*Assistant Professor (Sr.G), Department of Computer Science and Engineering, KPR Institute of Engineering and Technology, Coimbatore 641407, Tamil Nadu. Email: kiruthika.jk@kpriet.ac.in*
[5]*Assistant Professor, Department of Information Technology, Velammal College of Engineering and Technology, Madurai 625009, Tamil Nadu. Email: psaraswathimtech@gmail.com*
[6]*Assistant Professor, Department of Electrical and Electronics Engineering, Karpagam Institute of Technology, Coimbatore 641021, Tamil Nadu. Email: divya.eee@karpagamtech.ac.in*
** Corresponding Author Email: dhiyanu87@gmail.com*

shortcoming of these techniques is their inability to deal with data ambiguity. In the same way as human thought, fuzzy logic works on a similar principle. Hence, it can be utilized to handle data ambiguity. Fuzzy logic has been incredibly helpful in solving classification problems. It permits the recurrence of classes and helps overcome ambiguity. A fuzzy rule-based system can also be used with if-then rules to progress interpretation and provide more insight into Besides being able to assign objects to multiple classes with different membership levels, it also has the advantage of being able to assign them. Fuzzy logic is an accurate and valuable technique for early diabetes prediction.

In order to diagnose diabetes, the F-KMC-SVM uses a variety of factors including blood glucose levels, Body Mass Indexes (BMIs), skin thickness, diabetes family history, and age-related changes in blood sugar levels. Prediction. The overall system performance was evaluated using the diabetic dataset. The fuzzy-based F-KMC-SVM method has improved accuracy in predicting diabetes compared to previous methods the proposed method has very high classification accuracy, indicating its effectiveness in accurately predicting diabetes.

2. Related Work

Early diagnosis and prevention of diseases, such as diabetes, are crucial in medicine. ML algorithms can help diagnose diseases. However, modern lifestyles often include added sugar and fat, which increase diabetes risk. It is essential to recognize diabetes symptoms to stay informed and take precautions [1]. Unlike cloud computing, local fog nodes near the data source provide a faster and more secure option for storing and analyzing information [2]. To develop a DL approach to real-time DFU localization, current automated solutions are based on segmentation or classification. A database of 1775 images from DFU can be compiled to create robust deep-learning models. To ensure the dataset's accuracy, two clinical experts identify areas of focus for the DFU and use software to analyze it [3]. To determine smartphone-connected blood glucose level (BGL) measurement accuracy in diabetic patients (T2D) using an electrophysiological biosensor presented in Type 2. Moreover, ML techniques can be implemented to explore their potential [4].

Modern continuous glucose monitoring (CGM) devices promise to provide timely awareness of glycemic status in patients approaching critical hypoglycemia. The challenge is to identify patterns leading to dangerous situations with appropriate progressions. This is so that patients can make treatment decisions based on their predicted glucose concentration levels [5]. By utilizing transfer learning techniques, examining a Siamese-like convolutional neural network (CNN) model is possible. These models differ from previous studies and demonstrate that they can process binocular fundus images as input. Furthermore, it can help understand correlations between these images to facilitate accurate predictions [6]. DL techniques can remove reflex features in image processing and detect redundancy through specific steps. A practical approach is to devise a grayscale global color coincidence algorithm. This helps to smooth the image and enhance its general quality by normalizing the brightness [7].

For smart healthcare systems, it is a necessity to analyze patient data accurately and make predictions about diseases and treatments. Fatalities can be avoided to a large extent through timely referrals for medical attention and prompt awareness of emergency. Accurate predictive analysis of data using ML algorithms holds significant promise in the medical field. [8]. A capable solution called the Artificial Pancreas (AP) combines autonomous insulin management and glucose monitoring. However, progressive use of the AP requires reliable information, such as dietary carbohydrate intake for bolus organization [9]. The existence of outliers and missing values in diabetes datasets makes it difficult to predict diabetes with accuracy and reliability, as labeled data is limited and outliers and missing values are prevalent [10]. Type 1 diabetics can manage their blood sugar levels by exercising, which has not been widely reported. Hospitals use insulin infusions and extra carbohydrates (CHO) to prevent hypoglycemia. Insulin provides feedback and utilizes a proportional-derivative controller known as the Safety Assist Feedback Factor (SAFE) layer. Conversely, CHO is managed through a predictive and calibrated proportional-derivative controller [11].

The Health CPS is designed to cater to patients susceptible to or already suffering from Non-Communicable Diseases (NCDs). The system relies heavily on artificial intelligence (AI) technology to accurately detect NCD risks like heart disease and diabetes by utilizing an integrated approach [12]. By employing advanced feature extraction and machine learning techniques, a framework can anticipate HbA1c levels using blood glucose data gathered from continuous glucose monitoring (CGM) sensors, with a two- to three-month lead time [13].

Models of glucose levels predicting hypoglycemia, which utilize a glucose-insulin mixture (GIM) and consider inter- and intra-individual variability, enable the effective integration of various models with adjusted parameters [14]. Using machine learning techniques, it is possible to accurately determine future blood glucose levels. This can help prevent dangerous hyperglycemic states and optimize diabetes treatment. Many methods in the literature can create predictive models from historical data, reducing the need for extensive computer calibration [15].

Although many studies have been conducted on this subject, most have focused heavily on network design without considering the pathological relevance of lesions [16]. DSPN is an early indicator of the possibility of diabetic wounds and foot ulcers not healing as soon as they are laid on the skin. As a consequence of diabetes, it is one of the most common complications, leading to increased medical expenses and a reduced quality of life for the patient. These conditions may result in low blood pressure, infections, amputations, and death [17].

There is a distinct multifractal component to the retinal microvascular network, including certain general characteristics, defects, and a certain spectrum of functions within the network. Early detection of DR can prevent or delay vision loss [18]. Detecting small lesions using conventional CNNs can be challenging. The distribution pattern of DR data is unbalanced, leading to an overestimation of quality. Furthermore, these can have a significant impact on final quality performance [19]. Most current DR diagnostic systems lack consistent and fine-grained annotation of training data, resulting in satisfactory performance and interpretation for ophthalmologists [20].

2.1. Contribution of the Paper

- This novel main objective is early risk prediction of diabetes categorization like type and type 2 using a machine learning-based approach.
- Initially, we prepared the collected dataset using the Z-score normalization method. Then we chose optimal features for diabetes categorization with the Threshold Recursive Feature Elimination (TRFE) technique.
- Next, we use the fuzzy-based K-means clustering method, which groups the selected features based on weight values, calculates their centroid point, and then makes a diabetes diagnosis using fuzzy logic.
- Finally, we categorize diabetes patients based on type 1 and type 2 using Support Vector Machines (SVM). This method supports the maximum diabetes values in the decision-making dataset.

3. Proposed Methodology

This paper's main objective is to categorize the most promising results for the early prediction of diabetes patients based on type 1 and type 2. Figure 1 denotes the machine learning-based SVM architecture for diabetes-type prediction. This proposed system consists of four processes: i) pre-processing; ii) feature selection; iii) decision-making; and iv) classification.

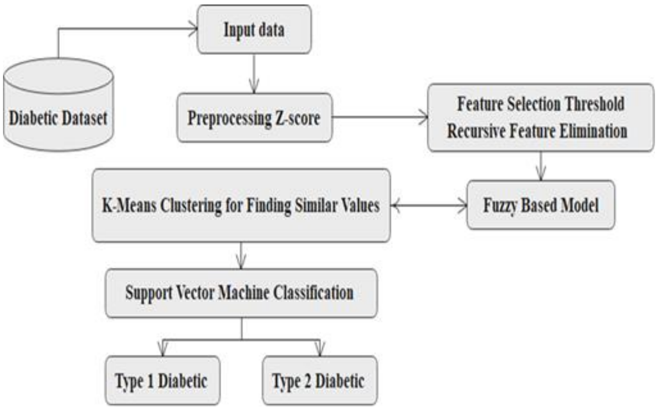

Fig. 1. Architecture for Categorization of Diabetes Patients

Firstly, we collect the dataset from the Kaggle repository called the Diabetes Prediction Dataset (DPD). Check for null values and irrelevant records, then select optimal features for diabetes patients. Next, we use a fuzzy-based model to make disease prediction decisions. Using the decision-making process, create a group for diseases, and then the proposed classifier identifies diabetes such as type 1 and type 2 in an efficient manner. Also, the following subsection describes how to predict diabetes.

3.1. Dataset Collection

This novel Diabetes Prediction Dataset (DPD) was collected from the Kaggle repository. This dataset includes the following descriptions: The number of instances is 100,000, and the number of attributes is 9. It incorporates gender, age, hypertension, heart disease, smoking history, body mass index, HbA1c level (glycated hemoglobin), blood glucose level, and outcome. This dataset contains 58552 females, 41430 males, and 18 others among 1000000 patients. Table 1 shows details,

Table 1. Characteristics of Diabetes Dataset

S. No	Parameters (Attributes)	Data description	Type of data
1	Total patients	100000	Integer
2	Gender	Male, female and others	Nvarchar
3	Age	Numeric	Integer
4	Hyper tension	0 and 1	Integer
5	Heart disease	0 and 1	Integer
6	Smoking history	Patient smoking history info	Nvarchar

7	BMI	weight in kg/(height in m)^2	Float
8	HbA1c_level	Numeric	Float
9	Blood glucose level	Numeric	Integer
10	Diabetes (outcome)	0 and 1	Integer

3.2. Preprocessing using Z-score

Data preprocessing is required to prepare the diabetes-type and DPD data to achieve high accuracy in disease prediction. Therefore, the proposed method uses the Z-score normalization method to handle inconsistent values and prepare the processed diabetes dataset. This method of data standardization is based on the mean and standard deviation of the original data.

$$S_i = \sum \frac{I_i - I_{mini}}{I_{maxi} - I_{mini}} \tag{1}$$

Missing values are analyzed in the dataset and estimated in equation, from that I_i refers to the diabetes input records of i^{th}each iteration I_{mini} and I_{maxi} denotes the minimum and maximum value records present in the dataset, respectively.

$$\mu = \sum_{i=1}^{j} \frac{S_i}{j} \tag{2}$$

The above is used to identify the mean (μ) values in the dataset. From there, j denotes the total number of records in the dataset.

$$\sigma = \sqrt{\sum \frac{(S_i - \mu)^2}{j}} \tag{3}$$

From equation 3, estimate the standard deviation (σ) values. Then we evaluate the Z-score normalization as shown in equation 4,

$$Z_{score} = \sum_{i=1}^{j} \frac{S_i - \mu}{\sigma} \tag{4}$$

In this experiment, the proposed Z-score method efficiently obtained the processed dataset from a collective one based on the mean and standard deviation.

3.3. Threshold Recursive Feature Elimination (TRFE)

The method used in this step is Threshold Recursive Feature Elimination (TRFE). The TRFE algorithm removes unnecessary or fragile features with no influence on the weights chosen for diabetes disease decision-making. This algorithm uses recursion to sort the attributes. This algorithm first generates a sample of all the features in the dataset. Sort the attributes according to their impact on weight values. Additionally, the TRFE algorithm removes features with the least correlation to the classification results and repeats the process to recalculate the attribute ranks.

Algorithm steps

Input:

Preprocessed dataset Z_{score}

Set of features $F_i = \{F_{1,} F_2 \dots F_n\}$

Output:

Important features ranking F_R

Begin

List of ranked features $\emptyset$

List of remaining features $\mathcal{R}$

Repeat for each i in $(1\ to\ n)$

Compute weight ($\mathbb{W}$) values of each feature

$$\mathbb{W} = \sum_{i=1}^{j} \gamma_i Z_{scorei} F_i$$

Estimate threshold ranking measures $R_K = \mathbb{W}^2$

Update the feature rank list U_{rank}

$$U_{rank} = \mathbb{W} + R_K$$

Remove smallest weight rank list feature

$$S_W = \mathbb{W} - R_K$$

End for each

Return important features ranking $F_R \leftarrow U_{rank}$

End

The above algorithm steps produce efficient feature selection for diabetes prediction based on the ranking method. In this experiment, the most significant features are analyzed, and the smallest ranking features are removed from the preprocessed dataset as a result. Let assume, F_i Denotes class label feature i ($F_i \in [-1, +1]$) and γ_i denotes Lagranigian multiplier.

3.4. Fuzzified Based K-means Clustering

At this stage, we used a Fuzzified based K-means clustering algorithm to make diabetes disease decisions based on feature selection attributes. In this algorithm, feature selection weight values are first grouped (cluster) based on weight. For each data point, computes the distance from the centroid point and assign the data point to the nearest cluster. Then based on fuzzy logic, the diagnosis of diabetes is made. From equation 5 estimate the closet distance in the cluster

$$Dis_{(a,b)} = argmin_b \left|\left|F_R - C_n\right|\right|^2 \tag{5}$$

where a, b are the two classes in the dataset; F_R refers to significant features, $Dis_{(a,b)}$ denotes the minimum distance between the data points and the centroid (C_n) in the cluster.

$$C_n' = \frac{\sum_{i=1}^{j}(F_R \in C_n)}{count\left[\sum_{i=1}^{j}(F_R \in C_n)\right]} \quad (6)$$

The above equation estimates the cluster's updated centroid point. The process ends when the sample points in each cluster stop changing, indicating that the j^{th}centroid has reached convergence. Decision-making is based on IF-THEN rules using fuzzy logic for diabetes. Here, some rules are presented

Rule 1:

IF patient has age > 40 AND BMI>60 THEN

IF (Sudden weight loss <δ AND
// δ threshold value

Days of slow wound healing< δ

Cholesterol>140)

THEN

Non diabetes patients

ELSE

High Diabetes

End IF

End IF

Rule 2:

IF Patient HbA1c_level >7.0 AND Insulin >250 AND

IF (Hungry> δ

Number of urinations per day> δ

Blurred vision<T_h)

THEN

Diabetes Medium

End IF

Rule 3:

IF patient smoking ==yes AND Blood glucose level>7.8

IF (check history with diabetic

Check heart disease

Check hypertension> δ

Thirsty> δ

Fatigue> δ)

THEN

Type 1 Diabetes

Else

Type 2 Diabetes

End IF

End IF

In addition to expert advice, medical parameters are also taken into consideration when developing fuzzy rules. Clinical datasets extracted from benchmark and streaming datasets will be considered as part of the final decisions, in accordance with expert recommendations and ambiguous time and interval constraints.

3.5. Support Vector Machine

At this stage, we used the Support Vector Machine (SVM) algorithm for classification based on the most insignificant classes of decision-making features. The SVM method creates a hyperplane in a multidimensional space. This hyperplane distinguishes between instances with various class labels. The set of data events used to construct the hyperplane is named a "support vector." The distance between the hyperplane and the nearest support vector has the highest diabetes categorization marginal value. Therefore, the proposed method identifies the maximum marginal values to predict diabetes categorization, like type 1 and type 2.

Algorithm steps

Input: Decision making rule

Output: Categorization of diabetes Type 1 or Type 2

Begin

Find the optimal values of the SVM parameters

Train the values using SVM model

$f \rightarrow f^*$;

While $f \geq 2$ do

$s_f \leftarrow$ Optimal values of the SVM parameters and dataset

$s_w \leftarrow$ Estimating the weight vector of $s_f(s_{w1} \ldots\ldots s_{w2})$

Ranking the features $\leftarrow w_{f1}^2 \ldots w_{fn}^2$

Minimum threshold values$\leftarrow$ training the lowest values

Remove minimum values from the dataset

$Rank_f$ ←Max.rank. dataset values

$f \leftarrow f - 1$

End while

$Marginal_1 \leftarrow values\ in\ dataset \in (Rank_2, \ldots Rank_{f^*})$

$$\text{Return}(Rank_1, \dots Rank_{f^*})$$

End

The above algorithm accurately categorizes diabetes types 1 and 2. This method classified diabetes disease types based on the most marginal (support) values.

4. Result and Discussion

The experiment used the Anaconda tool and Python as a portable, scalable, and interpretable programming language. System parameters used in the test: CPU at 3.30GHz, RAM at 16GB, SSD at 120GB, etc. Table 2 contains the parameter table.

Table 1. Simulation Parameter for Diabetes Disease Prediction

Parameter	Ratings
Tool used	Anaconda 2023.07-2
Language used	Python 3.11
Dataset name	Diabetes Prediction Dataset (DPD)
Number of records	100000
Training	83000
Testing	17000

4.1. Performance evaluation

This section evaluates the proposed comparison by considering parameters such as recall, false classification, accuracy, F-measure, and precision. The existing methods are Diabetic Sensorimotor Polyneuropathy (DSPN) [17] and Convolutional Neural Network (CNN) [19].

Table 3: Analysis of Precision Performance

Precision Performance in %			
Comparison Methods/Iterations	**CNN**	**DSPN**	**FKMC-SVM**
20	53.58	66.16	73.62
40	62.62	73.59	77.89
60	71.03	86.87	90.65
80	75.51	89.23	94.87

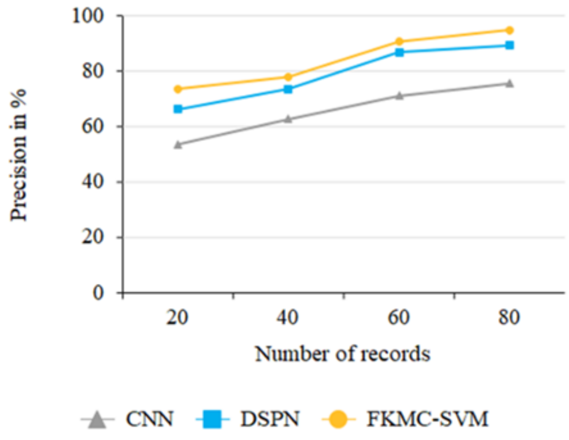

Fig. 2. Precision Performance

Figure 2 and Table 3 illustrate precision performance compared with other classifiers. The proposed method achieved 94.87% more precision performance for 80 iterations than other methods.

Table 3: Analysis of Precision Performance

Recall performance in %			
Comparison Methods/No of Iterations	**CNN**	**DSPN**	**FKMC-SVM**
20	51.05	65.09	71.54
40	60.35	71.25	76.95
60	70.55	85.98	89.21
80	73.61	85.35	93.29

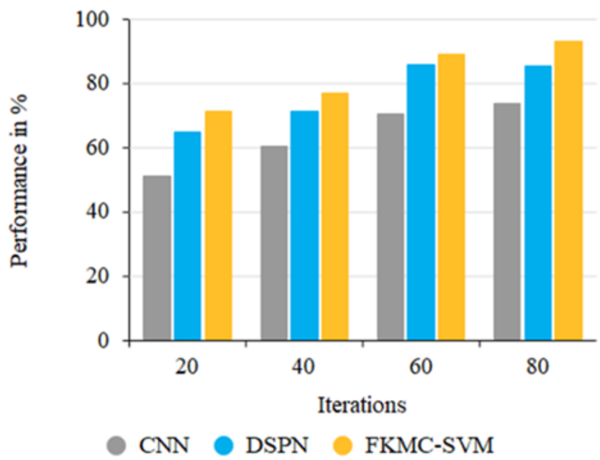

Fig. 3. Result of Recall Performance

Figure 3 and Table 4 depict recall outperforming different ML methods for diabetes prediction. The proposed FKMC-SVM algorithm has 93.29% recall performance; likewise, the existing algorithm DSPN has 85.35%, and CNN has 73.61% recall performance for 80 iterations.

Table 5: Analysis of F1-score

F1score Performance in %			
Comparison Methods/No of Iterations	**CNN**	**DSPN**	**FKMC-SVM**
20	56.05	67.68	74.26
40	63.16	74.65	80.15
60	72.98	87.03	93.25
80	76.48	88.21	95.41

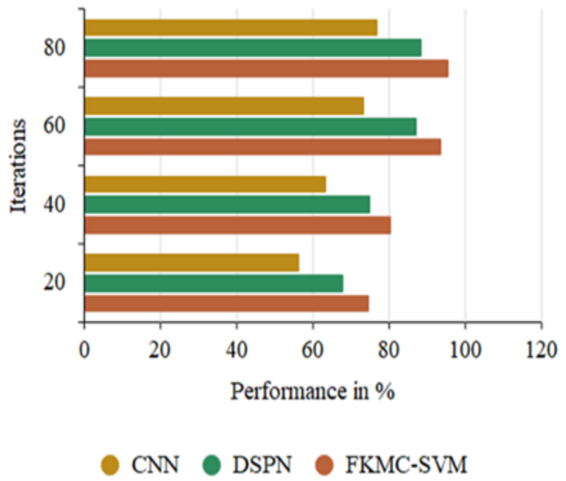

Fig. 4. Result of F1-score Performance

Furthermore, Table 5 and Figure 4 present F1-score performance with various classification methods. The proposed FKMC-SVM algorithm obtained 95.41% more F1-score to predict diabetes disease categorization than previous methods.

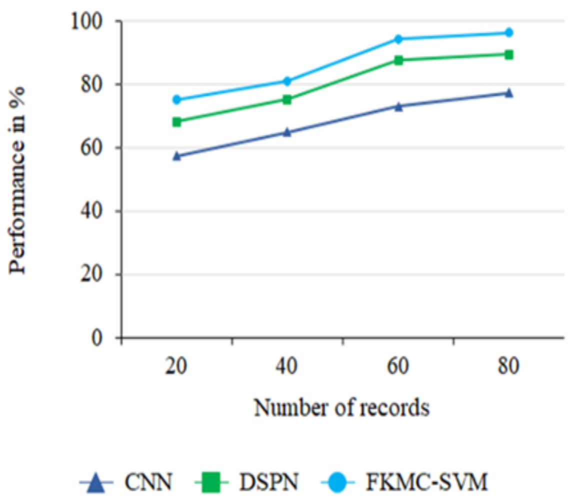

Fig. 5. Classification Performance by Risk Levels

Table 6: Result of Classification Performance for Analysis

Classification performance in %			
Comparison Methods/No of Iterations	**CNN**	**DSPN**	**FKMC-SVM**
20	57.35	68.36	75.16
40	64.85	75.35	81.09
60	73.14	87.68	94.35
80	77.25	89.57	96.24

Table 7: Analysis of False Classification Performance

False Classification Performance in %			
Comparison Methods/No of Iterations	**CNN**	**DSPN**	**FKMC-SVM**
20	15.98	12.44	10.23
40	25.42	19.84	13.48
60	35.05	23.59	18.06
80	40.55	27.21	21.03

Table 7 describes false classifications produced by various methods. For diabetes disease categorization prediction, the proposed FKMC-SVM algorithm has a 21.03% false-positive rate. The proposed method delivers less false-rate performance.

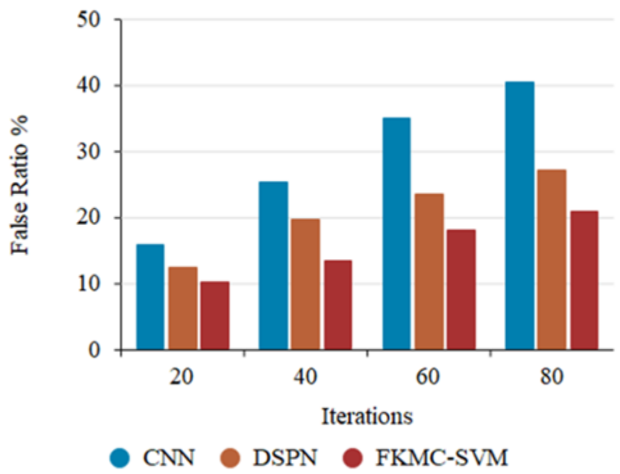

Fig. 6. Analysis of False Classification Ratio

The proposed method achieves less false rate than other methods presented in Figure 6. The proposed false rate is 21.03%; similarly, the existing method results are DSPN at 27.21% and CNN at 40.55%.

5. Conclusion

This paper presents fuzzy-based K-means clustering and the Support Vector Machine (FKM-SVM) method for diabetes disease categorization. Initially collected, the diabetic dataset is from the standard repository, and the second step is pre-processing to reduce the imbalanced data, normalizing the values from the dataset by Z-Score normalization, and then selecting the features depending on the margin values utilizing Threshold Recursive Feature Elimination (TRFE) to eliminate the values from the pre-processing dataset depending on the maximum threshold values of the recursive features in the dataset. Then the fuzzy-based method is used to decide diabetes using fuzzy logic to create interpretable models, to diagnose diabetes early based on these classifiers, KMC and SVM, and to design fuzzy rules. KMC refers to a collection of data points where points in one location share similarities or connections but differ from those in another cluster. Additionally, optimizing support vector machines with larger datasets may provide more accurate results and predict diabetes likelihood in both Type 1 and Type 2. The simulation result achieved a higher accuracy of 96.24% for disease prediction than other methods.

References

[1] U. Ahmed et al., "Prediction of Diabetes Empowered With Fused Machine Learning," IEEE Access, vol. 10, pp. 8529-8538, 2022.

[2] P. G. Shynu, V. G. Menon, R. L. Kumar, S. Kadry and Y. Nam, "Blockchain-Based Secure Healthcare Application for Diabetic-Cardio Disease Prediction in Fog Computing," IEEE Access, vol. 9, pp. 45706-45720, 2021.

[3] M. Goyal, N. D. Reeves, S. Rajbhandari and M. H. Yap, "Robust Methods for Real-Time Diabetic Foot Ulcer Detection and Localization on Mobile Devices," IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 4, pp. 1730-1741, July 2019.

[4] M. Z. M. Shamim, S. Alotaibi, H. S. Hussein, M. Farrag and M. Shiblee, "Diagnostic Accuracy of Smartphone-Connected Electrophysiological Biosensors for Prediction of Blood Glucose Level in a Type-2 Diabetic Patient Using Machine Learning: A Pilot Study," IEEE Embedded Systems Letters, vol. 14, no. 1, pp. 27-30, March 2022.

[5] M. Gadaleta, A. Facchinetti, E. Grisan and M. Rossi, "Prediction of Adverse Glycemic Events From Continuous Glucose Monitoring Signal," IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 2, pp. 650-659, March 2019.

[6] X. Zeng, H. Chen, Y. Luo and W. Ye, "Automated Diabetic Retinopathy Detection Based on Binocular Siamese-Like Convolutional Neural Network," IEEE Access, vol. 7, pp. 30744-30753, 2019.

[7] H. Kaushik, D. Singh, M. Kaur, H. Alshazly, A. Zaguia and H. Hamam, "Diabetic Retinopathy Diagnosis From Fundus Images Using Stacked Generalization of Deep Models," IEEE Access, vol. 9, pp. 108276-108292, 202.

[8] V. K. Daliya, T. K. Ramesh and S. -B. Ko, "An Optimised Multivariable Regression Model for Predictive Analysis of Diabetic Disease Progression," IEEE Access, vol. 9, pp. 99768-99780, 2021.

[9] G. Annuzzi et al., "Impact of Nutritional Factors in Blood Glucose Prediction in Type 1 Diabetes Through Machine Learning," IEEE Access, vol. 11, pp. 17104-17115, 2023.

[10] M. K. Hasan, M. A. Alam, D. Das, E. Hossain and M. Hasan, "Diabetes Prediction Using Ensembling of Different Machine Learning Classifiers," IEEE Access, vol. 8, pp. 76516-76531, 2020.

[11] Beneyto, A. Bertachi, J. Bondia and J. Vehi, "A New Blood Glucose Control Scheme for Unannounced Exercise in Type 1 Diabetic Subjects," IEEE Transactions on Control Systems Technology, vol. 28, no. 2, pp. 593-600, March 2020.

[12] R. Ferdousi, M. A. Hossain and A. E. Saddik, "Early-Stage Risk Prediction of Non-Communicable Disease Using Machine Learning in Health CPS," IEEE Access, vol. 9, pp. 96823-96837, 2021.

[13] M. S. Islam, M. K. Qaraqe, S. Belhaouari and G. Petrovski, "Long Term HbA1c Prediction Using Multi-Stage CGM Data Analysis," IEEE Sensors Journal, vol. 21, no. 13, pp. 15237-15247, 1 July1, 2021.

[14] W. Wang, S. Wang, X. Wang, D. Liu, Y. Geng and T. Wu, "A Glucose-Insulin Mixture Model and Application to Short-Term Hypoglycemia Prediction in the Night Time," IEEE Transactions on Biomedical Engineering, vol. 68, no. 3, pp. 834-845, 2021.

[15] Aliberti et al., "A Multi-Patient Data-Driven Approach to Blood Glucose Prediction," IEEE Access, vol. 7, pp. 69311-69325, 2019.

[16] S. Huang, J. Li, Y. Xiao, N. Shen and T. Xu, "RTNet: Relation Transformer Network for Diabetic Retinopathy Multi-Lesion Segmentation," in IEEE Transactions on Medical Imaging, vol. 41, no. 6, pp. 1596-1607, 2022.

[17] F. Haque, M. B. I. Reaz, M. E. H. Chowdhury, F. H. Hashim, N. Arsad and S. H. M. Ali, "Diabetic Sensorimotor Polyneuropathy Severity Classification

Using Adaptive Neuro Fuzzy Inference System," IEEE Access, vol. 9, pp. 7618-7631, 2021.

[18] M. M. Abdelsalam and M. A. Zahran, "A Novel Approach of Diabetic Retinopathy Early Detection Based on Multifractal Geometry Analysis for OCTA Macular Images Using Support Vector Machine," IEEE Access, vol. 9, pp. 22844-22858, 2021.

[19] He, T. Li, N. Li, K. Wang and H. Fu, "CABNet: Category Attention Block for Imbalanced Diabetic Retinopathy Grading," IEEE Transactions on Medical Imaging, vol. 40, no. 1, pp. 143-153, Jan. 2021.

[20] Y. Zhou, B. Wang, L. Huang, S. Cui and L. Shao, "A Benchmark for Studying Diabetic Retinopathy: Segmentation, Grading, and Transferability," IEEE Transactions on Medical Imaging, vol. 40, no. 3, pp. 818-828, March 2021.

[21] Gothane, Suwarna, K. Srujan Raju, Saikumar Tara, and P. Devika. "Unsupervised learning towards hierarchical clustering." In AIP Conference Proceedings, vol. 2477, no. 1. AIP Publishing, 2023.

[22] J. Narasimharao, A. V. Reddy, R. Regulagadda, P. Sruthi, V. Venkataiah and R. Suhasini, "Analysis on Rising the Life Span of Node in Wireless Sensor Networks Using Low Energy Adaptive Hierarchy Clustering Protocol," 2023 IEEE 5th International Conference on Cybernetics, Cognition and Machine Learning Applications (ICCCMLA), Hamburg, Germany, 2023, pp. 651-659, doi: 10.1109/ICCCMLA58983.2023.10346731.

[23] Madala, S., Suvarna Vani, K., Jalapally, P., Srujan Raju, K., Somasekhar, G. (2022). Framework for Diabetic Retinopathy Classification. In: Satapathy, S.C., Bhateja, V., Favorskaya, M.N., Adilakshmi, T. (eds) Smart Intelligent Computing and Applications, Volume 2. Smart Innovation, Systems and Technologies, vol 283. Springer, Singapore. https://doi.org/10.1007/978-981-16-9705-0_58