ISSN: 0974-5823 Vol. 6 No. 3 December, 2021

International Journal of Mechanical Engineering

A Premature Diabetic Prediction Using Optimal Firefly Feature Selection Based on Angular Vector Matrix Transformation Neural Classification

F.Regina Ananthi[1],Dr.Renuga Devi[2]

[1]Research Scholar, Vels Institute of Science, Technology and Advanced Studies (VISTAS), Chennai,
[2]Assistant Professor, Department of Computer Science,Vels Institute of Science, Technology and Advanced Studies(VISTAS), Chennai.

Abstract

Today, diabetes is one of the most common, chronic, and the deadliest diseases in the world. Thus, early diagnosis and treatment are required to prevent diabetes and its associated health problems. The early detection of diabetes is very important for timely treatment since the progression of the disease can be stopped subsequently.The occupancies evolved from diabetic data create problematic features whichlead irregular weights to identify the importance of classification. To resolve this problem, improved Subset Feature Selection BasedOnAngular Vector Matrix Transformation Neural Network Algorithm (AVMT-NN) is proposed to predict the occurrence of diabetes in the future determination with optimal firefly algorithm.The first stage preprocesses the input data set to remove the noisy features and the firefly is optimized with Disease Influence Measure (DIM) for predicting spectral threshold values to create subset features. Once the model is trained to predict the subset features, then individuals can self-assess the risk of diabetes based on relational feature (RF) weights. Towards diabetic prediction, in the second stage, the classification predicts the risk by class that depends on the RF value and Angular Vector Matrix Transformation Neural Network creates the multiple neurons modulated searches. The method performs diabetic prediction based on the risk by class. Different approaches to disease prediction have been considered and their performances in disease prediction has been compared to other methods.

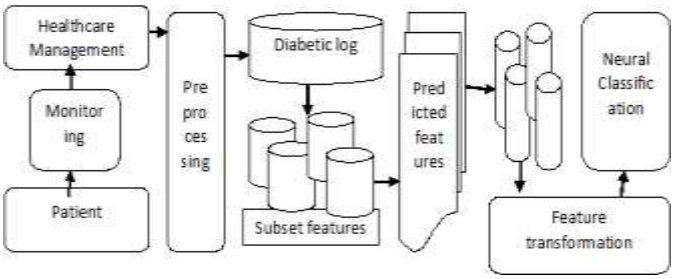

Figure 1 Diabetic feature analysis and transformation.

However, the fundamental standards and procedures utilized for information mining in diabetic information bases may vary for feature selection and classification approaches. Figure 1 shows Diabetic feature analysis and transformation in diabetic patient records. Information mining is one of the subset features for process in diabetic forms. The general objective of the information mining process is to remove data from an informational index and transform it into a logical structure for additional utilization.This procedure has turned into an inexorably unavoidable factor in every aspect of medicinal science examination. Information mining issues are regularly explained using diverse methodologies from both PC sciences, for example, multi-dimensional databases, machine adapting, delicate processing and information perception; and insights, including speculation testing, grouping, classification, and relapse strategies.

Keywords: diabetic prediction, neural network, premature treatment, feature selection, influence measure.

1. Introduction

In healthcare, implementing systems that use computing to improve quality of life and improve health as health services are found. Because of the decline in birth rates and the demographics of the aging community, the number of elderly and chronic patients in the family, outpatient health management needs to be increased. As a result, the need for ubiquitous health (U-health) services has increased. Diabetes is caused by a lack of production of insulin. Age, family history, weight, and previous gestational diabetes are some of the factors that facilitate the manifestation of diabetes. Type 2 diabetes (T2DM) is caused due to insufficient production of insulin hormone. Diabetes is divided into three categories: Type 1 Diabetes Mellitus (T1DM) can be diagnosed in with gestational features and the result are verified with sequence feature closest to diabetic fetures. Lifestyle modification or medication and injection or treatment are to be recommended based on the needs of diabetics. This treatment is very important to prevent dangerous complications. In addition, patients with diabetes need a blood glucose meter to measure their serum blood sugar levels daily. In addition, they need to analyze number of feature to predict he disease levels. In many cases, the symptoms of hyperglycemia in the blood are not immediately apparent. If not properly managed, problems such as cataracts, vascular and kidney problems, disorders of the nervous system, and conjunctive loss can occur.

Therefore, diabetics need to monitor and control their diet, exercise therapy, medication, etc., and constantly manage their blood sugar. However, diabetics find this difficult to manage on their own. They need constant management support by analyzing the diabetes which is important in predicting early results using the data mining approach. The purpose of this research is to evaluate the effect of diabetic patients for premature treatment using feature selection and

classification approach on different stages in a recommended and random group for diabetics in the early stages of the disease. By considering the proposed implementation. An Optimized Spectral Neural Classification Model based on Multiple Kernel Fork Joining Clustering (SNC-MKFJC) has intent a classical approach in diabetic prediction followed to improve the feature selection and classification approach.

This implementation integrates the Feature analysis importance to realize the potential aspects weights for diabetics through diagnostic and diagnostic tools with minimal variance performance definitions. The importance of feature selection is based on the subset evaluation such as ensuring adequate control of dimension problem for predicting the diabetic classes. This task is to determine the classification approachusing artificial neural network (ANN). so that minimum number of frequency tracking can be taken to deliver the patient's diabetic future level with a good understanding of the model to make classification accuracy. The influence of a unit depends on the weighting; where the input signal of neurons meets is called the synapse. ANN works for both supervised and unsupervised learning techniques. Supervised learning was used in our study because the output is given to the model. In supervised learning, both input and output are known.The classification supports high prediction by the risk of evaluation in the diabetic dataset to categorize class by reference. This highly improves the classification accuracy well to recommend for diabetic early treatment.

2. Literature survey

The Adaptive Model Predictive Control (MPC) method is used to predict diabetic features observed in patients with type 1 diabetes mellitus with closed dual hormone synthetic pancreatic function [1]. Dual hormone AP measurement uses continuous subcutaneous glucose monitoring and manages glucose whichare the important features having concentrations into the subcutaneous screen of rapidly acting insulintype based feature analysis [2] [3]. The naive Bayes model handles the features from the linear parameter variable which consists of a multi-objective classification supportive and is designed for hypotonic, hyperactive, and normal blood sugar conditions [4].

The key feature is that the controller is a stacked one on one framework that guarantees stability and performancefor finding diabetic-based feature analysis [5]. In the balance between postprandial hyperglycemia and hypoglycemia: Type 1 diabetes has autoimmune blood sugar. The non-linear mismatch controller design makes a trade-off between the elimination of hypoglycemic events and the limitation of post-meal hyperglycemia.

Multiple Model Strategy an alternative closed-loop approach to subcutaneous insulin delivery for type 1 diabetes was evaluated [6]. The glucose-insulin regulatory system has been considered as a various operating point modeling system around the anatomical characteristic. Graphics in cloud computing services such as a new Geometric processing unit (GPU), and the largest parallel computing can be used in medical data analysis applications [7] [8]. In addition, workshops suitable for deep learning computing are usually a large set of data that can be distributed between multiple units per GPU.

Multiple feature units are available to allow parallel classification in an artificial neural network [9]. Type 2 assesses non-hospital costs including short-term and long-term hospitalization for diabetic typing-related complications [10].The prevalence of type 2 diabetes (T2DM) is determined in the new cases based on managing resource data called management information [11]. Diabetes is one of the biggest population well-being problems in the world. Clinical diagnostic programs show that diabetes and blood pressure have their predictive mechanisms of action while monitoring the parametersthat have become the most important way to control glucose levels [12].

The predictive performance of a machine learning decisive neural network (Dec-NN) system depends entirely on the effective selection of components [13]. One of the keys to the selection of facilities is statistical models and machine learning challenge tasks [14]. The high-efficiency feature is selected to use the Least Absolute Shrinkage and Selection Operator (LASSO) method to predict diabetes. The Bayesian Network (BN) can predict the spread of diabetes from the natural relationship between environmental risk factors used in Naïve Bayes [15]. By using the BN model, we can capture the correlation between direct and indirect risk levels.

Diabetes mellitus (DM) is defined as a group of metabolic disorders that exert significant global pressure on human health. Aspects of diabetes have led to the creation of a large body of data. In the health sector, diabetes is confusing as far as data is concerned (unstructured). Therefore, the system needs to emphasize these large-scale diabetes data for better and more accurate prediction results in support vector machine (SVM) [16]. The main task of this forecast is to select features related to the forecast. With the help of feature selection algorithms, a good prognosis for diabetes can be achieved [17] [18].

The main task of this forecast is to select features related to the predicting and feature-based classification methods. With the help of feature selection algorithms, a good prognosis for diabetes can be achievedusing support vector machines but least level is achieved in classificationaccuracy [19]. Monitoring Blood glucose without current blood glucose monitoring devices causes pain and discomfort due to prolonged diabetes. Frequent monitoring of glucose is an important part of diabetes management. According to reports, diabetics include abnormally low levels of acetone concentration in the breath and a gradual increase in the patient's blood glucose level [20]. The concentration of certain biomarkers into blood glucose levels is associated with diabetes mellitus (BGLs). Therefore, it is essential to screen and predict BGLs in diabetics by examining human respiration [21]. Classification of diabetic features based on documentary data is one of the computational diagnostic methods.

Cardiac autonomic neuropathy (CAN), a primary complication of diabetes, is characterized by a gradual increase in autonomic nerve fiber damage that causes abnormal dynamics of the ventricular repolarization (VR) and sub-clinical condition Positive treatment results to be diagnosed are rarely the clinical condition [22]. Due to significant changes in high-risk factors for diabetics, many patients have complications and avoidable injuries [23]. High-

risk factors can help reduce the incidence of enhanced identification complications.

The most confirmed case associated with a sudden increase and changes in human life is type 2 diabetes. The amount of sugar in the blood is higher than normal levels is a condition [24]. Classification systems are widely predictive models to analyze patient data or rules which have been used in the medical field to shoot. To provide a better classification of diabetes [25]. Several existing methods are being used to classify and implement diabetes data sets.

3. Materials and methods

In this proposed system, Angular Vector Matrix Transformation Neural Network Algorithm (AVMT-NN) is proposed to predict the occurrence of diabetes in the future determination with an optimal firefly algorithm. This initiates the subset feature selection-based classification using ANN.

Initially, the method preprocesses the input data set which removes the noisy features. The Disease Influence Measure (DIM) was estimatedto find the range of the diabetic margins. Once the model is trained to predict the subset features, the DIM contains the disease influence margins with relative features to construct patterns. Depending on the disease pattern, feature margins levels are verified with the risk of the margin with the relational feature (RF) weightage. Based on the RF value, Angular Vector Matrix Transformation Neural Network creates the multiple neurons modulated searches thereby the method performs diabetic prediction.

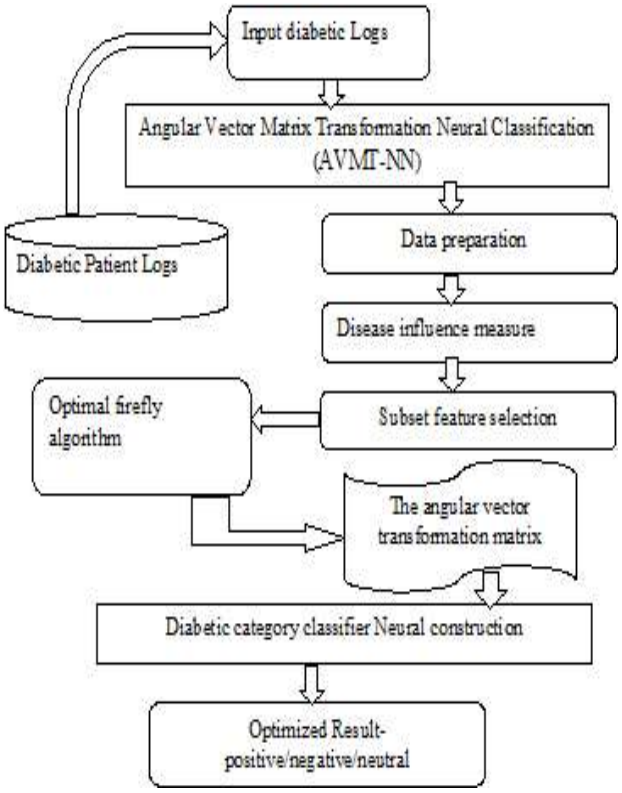

Figure 2 Architecture diagram for proposed AVMT-NN

Similarly, the ANN is composed of multiple nodes that are connected to each other. The connection between units is represented by a weight. Figure 2 shows the architecture diagram for the proposed system AVMT-NN. The objective of ANN is to convert input into significant output. Input is the combination of a set of input values that are associated with the weight vector, where the weight can be negative or positive. The hidden layer output contains a decisive tree logical function unit to train features effectively which depends on feed weight. The logical decision functions split the neural decision closest to the spectral threshold for detecting complex patterns and learning feature support weight which is a function that sums the weight and maps the result to the output.

3.1 Data preparation

Diabetic data analysis contains a high volume of data in the form of high dimensions to point out various attributes. Due to the collection of vast information, the situation is very noisy. (I.e., irrelevant information, data mismatch, empty fields, other irrigated signs. etc.). The collection of features contains the attribute which is in the redundant form to select the Diabetic patterns**.** The logs are collected under different levels of diabetic patients type 1. In this stage, the method first reads the logs available in each class of diabetic features. Then the list of features available in the trace is identified from the patient's observations. Using the feature list identified, the method verifies the availability of features and range of the value for all the features present in the feature list. If any of the traces identified as incomplete or having missing values, then the method eliminates the trace to continue cancer detection and analysis.

Algorithm: Pre-processing

Input: Initialize Collective diabetic dataset- Cds

Output: Filter Processed dataset -Fds.

Step 1: Cds= observe collective records labels.

Step 2: compute For. (Cds→I at the initialization at J feature)

Step 3: check if empty→true and its Range

Remove records, check cleansing null attribute, stemming the progress

Step 4: Fds= return rearranged data records

Step 5: end if end for;

Step 6: return as redundant R-Fds

The above algorithm describes diabetic patient records and the pre-processing stage to reduce the dimension based on an attribute by filtering the values. All the records contain multiple attributes as features which are referred to the single patient information. The log would contain numerous entries, among them there would be the number of incomplete records which contain missing values. Such missing values are identified and the records with incomplete features are eliminated from the set. Such noise-removed data set is used to perform dimension reduction to make Disease influence analysis.

3.2 Disease influence measure (DIM)

In this stage, Disease influence measures (DIM) are estimated based on the non-homogenous value of Diabetic data for finding disease infection margins. In this process, the individual features fromDiabeticthreshold values which are relatively identified by links of connectivity between the communal features, for example, Glucose level range are normalized or high which is relatively close to insulin which is estimated because the frequent Diabetics are associated

based on repeated count relative feature equalized to get the influence rate. These Diabetic variables have identity links to the clearance of the relational connectivity between the Diabeticfeatures. This reduces the non-relational terms as well as the cleansing way and combines distortional features same as attribute relational variables to reduce the non-relational terms.

Input: preprocessed dataset Ps

Output: Relative weightage Rds set

Start

Step 1: Initialize the processed set Ps

For each feature (read ←Diabeticid)

Step 2: For each attribute Ca from Ps

Identify the threshold closest terms Ct value

Max term value (max variable count → confidence)

Rearrange the Diabetic Id

Create links between the term Diabetic id→cd1,cd2.

Step 3: compute the key feature relationkt

For each Diabeticfeaturekt from Ps

For each Diabetic term Kt from Ps

If Kt € max term then

Identify max count relation with other attributes.

Relation Set Rs = ∑(Concepts€Kt)+Ps.

Compute the Number of attribute relations it has.

Kts = ∑ Relations €Gi similar term

Compute the max count attribute value

mval = ∑feature Links(Kt)<-∑mval(Kt)

Create link Kt identified relative feature links

Relative link Rt←kt+ca;

End

End

End

Step 5: compute the conceptual feature links→CRL.

For each feature←RL

CRL = = ∑Concept (Links (Kt))) € ∑Concept (Ca)!=Ps

Compute relative feature→RLK

RLK = $(Kt + Ca) + NIL$

Add to feature link set CRL←RLK;

End

End

Stop.

The above algorithm reduces the frequent intervals between the connected attribute links that are related to each other. The maximal Diabetic has count which refers to the repeated Diabetics. So, the distinct frequency will be evaluated to originate the data.

3.3 Feature vector decomposition

The feature weight similarity based on Singular Vector Decomposition (SVD) is used to find the relative closeness of the Diabetic pattern. The Diabetic key feature terms are analyzed with an extracted related terms from the disease influence measure to find the direction of influence occurrence ofDiabetic objectives. Based on consecutive features, the vector decomposition constructs a pattern under different level based on influence measure. These analyses of feature values relation between current and future increase values based on the frequency level of patterns.

Algorithm:

Input: frequent processed featureDiabetic set:

Output: optimized relational class

Step1: Compute the frequent Diabetic set processed features

Start

Observe the term Diabeticfeature set Ts.

Feature weightterm observation Ts→O.

Step 2: Compute the Diabetic set

For each Diabeticfeature weight index Ti from Ts

Observe bound measure in feature weight relation SMr = Nc/Tn.

Non represented class→Nc = Ti representation of the number of class T

The term point of Diabetic set →Tn-closure weightage representation of terms

End observation

Step 3: compute the category of feature weight observation

Representing For each class C

Process the closeness measure by feature weight relation Scm =$\int \frac{sbm}{Numberoftermspresentinotherclass}$

Step 4.Relational process can be identified by the following equation,

RT (t) =$\frac{term\ to\ searching\ the\ data\ in\ attribute}{total\ number\ of atrributr\ in\ the\ dataset}$,*(t→ in terms of time-frequency)

Step 5: Defending to Diabetic set identification feature set

Relational feature weight level SRT=$\frac{\log(\sum attribute\ \ relation\ text(RT))}{total\ attribute}$

Step 6: feature weight relational identification from feature weight measure.

RI=$\frac{\mathrm{RT(t)}\times SRT}{Total\ time\ taken}$

End

Select the closure measure at top rate class C = O(Max(Scm))

Step 7: Compute the spectral marginal relations

$\sum$ (Vi, Vj) =. $\frac{(distance(\text{Vi,Vj})+\beta(Seq)*\mu*(d(Vi)+d(Vj))}{N(Vi,Vj)*2*seq*\max(d(Vi)-d(Vj))}$

Step 8: Non elements are observed Ne = $\sum Terms(Ts) \nexists O(c)$

Stop

Where d (vi) and d (vij) are to reveal the row of the decision node of the tree with Vi and Vij in the singular node respectively. The distance (Vi, Vj) is the V term. And the number of margins on the narrow path between the Vij is Vi and the Vij is also the weight value of all the edges on the short track between my (Vi, Vij). Seq is the maximum depth of the tree with the entity, β is an adjustable parameter. Based on the relational term extraction the non-relative feature elements be avoided to select the feature terms.

3.4 Firefly optimization for subset feature selection

In this stage, fireflies have properties that can be extracted from peak ratings. Each firefly is assigned by light intensity and total extraction features and the clear features are being the best selected. This high boundary area is drawn towards the center and defines the definitions that are based on the properties of particles of similar classes with properties extracted from the random area.

Let us consider Fv as the feature vector values pointed on the feature matrix. On selecting a training feature, Define m, $nando$with some random values (here 0.2, l.0, and l.0 are considered respectively).

Let $X = Xi$ (i = 1,2,3 . . . . . z)

Where 'n' is the number of particles and 'X' is the population of fireflies.

$DefineI = rand(Fv)$..(2)

Where I is the light intensity. Updating the observation coefficient as

$mi = pi - pj\ 2 + qi - qj\ 2\ m_{i=}\sqrt{(p_i + p_j)^2} + (q_i + q_j)^2$

(3) Where i=I,2, 3, ... z, j=I,2, 3, ... m. Final updates are expressed as

$pn = pni \times 1 - n + pnj \times n + (\text{rand} - 0.5)$

$qn = qni \times 1 - n + qnj \times n + (\text{rand} - 0.5)$

When the light intensity gets updated after some iteration, the final values are indicated as $fnt = I\ (x, y)$ Exact fitness value bv = min(fnt) Exact best fitness value

$Tf = Fv\ (bv)$ Selected best feature and create subset clusters.

Random feature values are formed based on functional types, and such particles of nature follow their region. All particles are concentrated in some region and the best functions for classifying these diseases are defined. Therefore, firefly optimization will work on the purpose of feature reduction techniques that ignore others that take into account similar features.

3.5Angular vector-matrix transformation neural classifier

In this stage, the matrix is construed to combine subset equivalence value for repetitive terms that can easily create diabetic occurrence depending on the pattern feature observation.The relativity depends on the number of times that the key terms have occurred. Once again, key terms also add up the most credible value for the unnecessary Diabetic events. The number of occurrences forming the feature relationships is identified using the frequency phenomenon of transformation feature weights. Using the feature weights, the patterns are generated in each transaction by computing max-min vector values at each feature observation.

Algorithm:

Input: Diabetic frequent set Pi, Pattern set Ops.

Output: Diabetic terms Pi.

Step1: initialize PI to null.

Step2: for each item I from SI

Identify the patterns where i1,i2.. present.

PS = number of Diabetic transactions

Compute for Process state Ps

Process interims tags of P→term (i) = max (count$^{\sum TS(i) == Pi}$)

End

Step 3 Compute support = Count/size (Ts).

Add to Min Max count MM = $\sum MM(i) + Support$

If count →(mean) then

Add I to frequent Diabetic set PI = $\sum PI(i) + I$

End.

End.

Step 4: stop.

The Diabetic factor depends on the key terms from the attribute value measure by repeat count terms. To find the frequency of repeat terms, the feature linkage utilizes the total Diabetic occurrence by the number of different Diabetics involved in various levels. For example, the different Diabeticfeatures represent a single relational pattern term which has involved in multi-attribute reference terms

3.6Optimized angular Neural Classifier

This angular neural network is constructed to find the risk of the diabetic's category based on the decisive rule constructed from ANN.This neural network has the ability to perceive both immediate and nonlinear illustrations that exist among data factors. When computing it, the input layer, middle layer, and output layer need to be selected. Here, an effective data classification neural network is used.

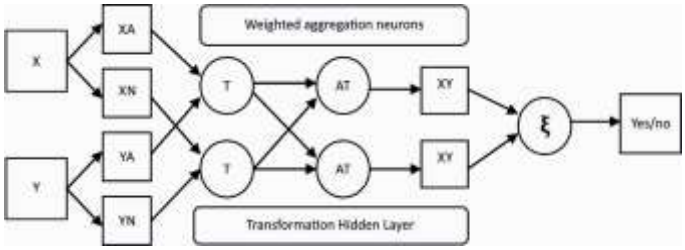

Figure 3 angular vector-matrix transformation classifier model

Each layer contains a log section called the source of data, called a weighted aggregation neuron, and after a short time,

all non-linear transitions are ended. Neurons interact with flexible weights with different layers of location. Depending on the size of the hidden layers and the size of the neurons in each layer, subsequent input layers, feed forwarddynamic four to six hidden layers, and 3 output layer are used.

Algorithm: Optimized angular Neural Classifier

Input: subset feature set Sfs-diabetic class threshold

Output: class by risk labels

Step1: Construct a Neuralclassifier on Threshold Neural nodes

For i=Sfs←Fs(feature set)

Attain input class for each features set Fs

Step 2: Compute the decisive neurons At Max function

Step 3: Train tested values comparison with each neural subset values Step2

Step 4: Split centroid closest and group the decisive nodesfor each class Cls

Cls→Check (Sfs→ max state) if max feed to next layer

Else

Return previous centroid decisive nodes Prs→back propagate (Bps)

Continue step 2; return Centroid class (Cnc)

Cnc→Sfs 1, 2, 3.. Update class on each class

Step 6: Each Feature calculate the fx value of the input layer

Step5: For each class Fs from Max Classes

Step6: compute the hidden feed layer at sub spectral points

For selective points Sp→Update each class F(x)

Each decisive class Dcs←Max(Sfs)

Split the class risk by class

Step 7: Return Cls.

To get to the diabetic feature processed with testing and training level, the neural networks are just promising, with no known answer. The framework should create a graph of the dataset by tracking sufficient binding weights. The amount of nodes in the input or data layer is set to select features ofdiabetic dataset which is trained with decisive logical rules at the rate of margin events. Diabetic data is set with 4 hidden nodes that are likely to have 4 sub-set classes which are trained at the event on running feature evaluations. The feature cluster groups are acquainted with a decision system produced by classified classes. The decision table is used to isolate space findings of insights regarding dependence factors by utilizing the possibility of values in diabetic sets. The dependence parts of the features is identified with the relational classes, and the type of them are instated as Link weights between input terminals, hidden layers and output layers to separately produce the result and support for premature treatment as well.

4. Result and Discussion

The diabetic prediction results verified under the big data forums using he machine learning objectives to process the diabetic data logs. The confusion matrix was generated to analyses he results. The proposed AVMT-NN produce higher performance level by testing the accuracy in sensitivity and specificity rate with false rate analysis and time complexity. The proposed method comparison are made with equivalent methods that are shown in table. The features are relatively selected from electronic health records collected from hospitalized dates I diabetics type 1 level.

Table 1 Processing tools and its variables

Processing tool	**Variables and its values**
Framework tool/language	Anaconda –Jupiter notebook/ python
Dataset used	Diabetic –Type 1 kaggle dataset
Number of features used	Diabetic labels <30
Testing progress	Confusion matrix
Type prediction	High, medium, low

The above table 1 listed the resultant parameters and values processed under i3 Intel configuration with 4 Gb of ram in Microsoft platform. The estimations of true values and positive values are trained with tested dataset. The classification accuracy was calculated using,

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

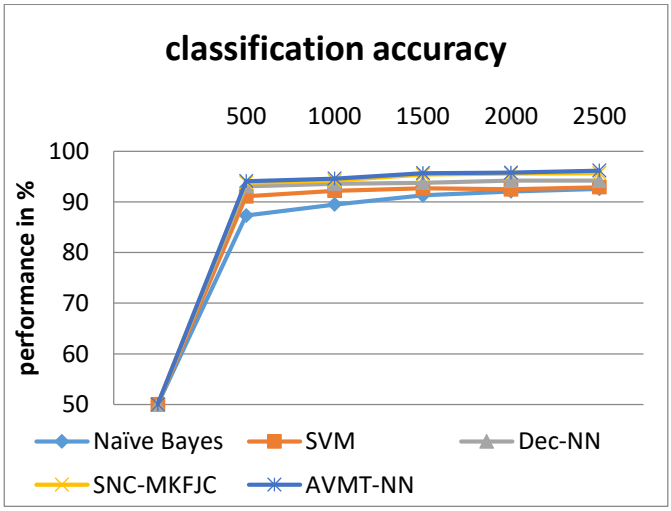

Figure 4 Impact of classification accuracy

The proposed system predicts the classes under different levels of class labels based on classification of risk evaluation. Figure 4 shows the classification accuracy produced by different methods as well the proposed system has high performance in classification accuracy compared to the other systems

Table 5 Impact of classification accuracy

Methods/ Datasets	Impact of Classification Accuracy in %				
	Naïve Bayes	SVM	Dec-NN	SNC-MKFJC	AVMT-NN
500	87.3	91.1	93.1	93.8	94.1

1000	89.5	92.2	93.6	94.2	94.6
1500	91.3	92.7	93.8	95.4	95.7
2000	92.1	92.5	94.2	95.6	95.8
2500	92.6	92.9	94.2	95.7	96.2

Classification accuracy defines the importance of prediction rate produced by different methods The naïve Bayes classification produce 87.3 %, support vector machine produce 91.1 %, decentralized neural network produce 93.1%, SNC-MKFJC produce 93.8 % and proposed AVMT-NN produce 94.1 % higher classification performance than other dissimilar methods. The sensitivity rate predicts the true positive values by the average mean rate of false negative rate observed form the evaluation.

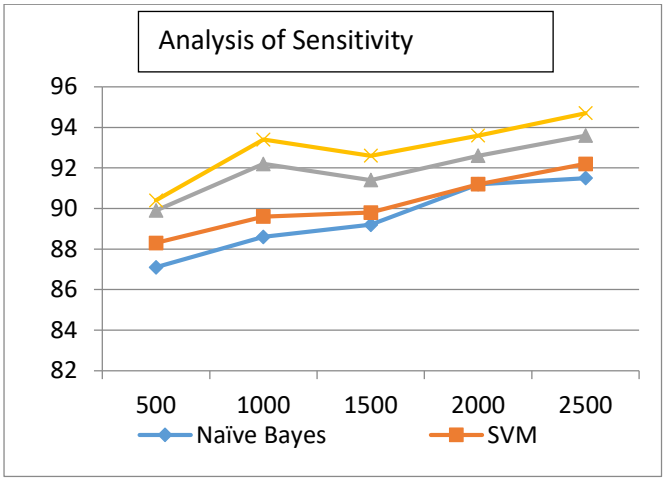

Figure 5 Impact of sensitivity analysis

The AVMT-NN esteem, produces 91.6% sensitivity rate as shown in figure 5, decentralized neural network accomplishes 88.3 % sensitivity yet SVM accomplishes only 87.1 % sensitivity. The sensitivity rate is calculated using , $\text{Sensitivity} = \frac{\text{TP}}{\text{TP+FN}}$. The proposed system has higher prediction in the sensitivity supportive measure to the classification accuracy.

Table 6 Impact of sensitivity analysis

Methods/ Datasets	Impact of Sensitivity Analysis in %				
	Naïve Bayes	SVM	Dec-NN	SNC-MKFJC	AVMT-NN
500	87.1	88.3	89.9	90.4	91.6
1000	88.6	89.6	92.2	93.4	94.3
1500	89.2	89.8	91.4	92.6	94.7
2000	91.2	91.2	92.6	93.6	95.2
2500	91.5	92.2	93.6	94.7	95.7

The table 6 shows the sensitivity performance compared to the other methods. The intensive rate defines the true negative values at the rate of averaged by false positive values.The specificity is calculated by. $\text{Specificity} = \frac{\text{TN}}{\text{TN+FP}}$

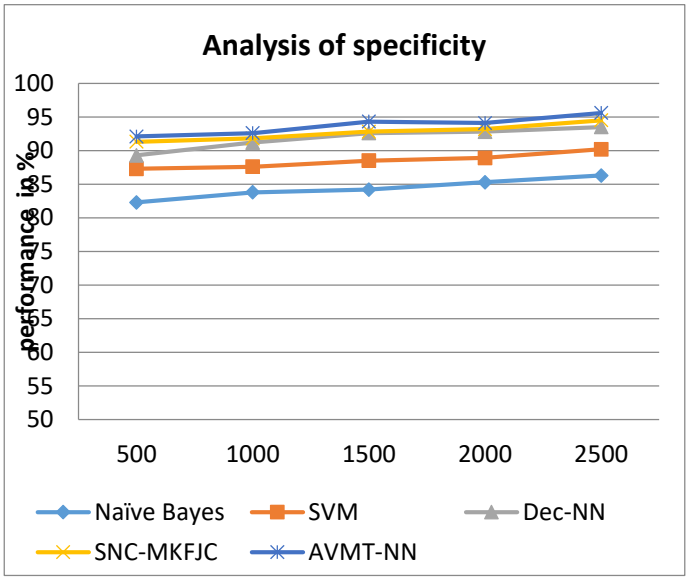

Figure 6 Impact of specificity

The contrast of specificity is shown in figure 6 with comparison of methods. The proposed AVMT-NN produce high performance compared to the other dissimilar methods. This achieves high specificity rate in diabetic classification.

Table 7 Impact of specificity

Methods/ Datasets	Impact of Specificity in %				
	Naïve Bayes	SVM	Dec-NN	SNC-MKFJC	AVMT-NN
500	82.3	87.3	89.3	91.3	92.1
1000	83.8	87.6	91.2	91.8	92.6
1500	84.2	88.5	92.6	92.8	94.3
2000	85.3	88.9	92.8	93.2	94.1
2500	86.3	90.2	93.5	94.5	95.6

The False rate defines the error congestion in worst-case measure from frequent analysis. The performance of the classification is low in this proposed system as well than other existing methods. The false extraction is calculated by False Extraction Ratio (Fer) = $\sum_{k=0}^{k=n} \times \frac{TotalDataset\ FailedtoClassify\ (Fer)}{Totalnoof Data(Fr)}$

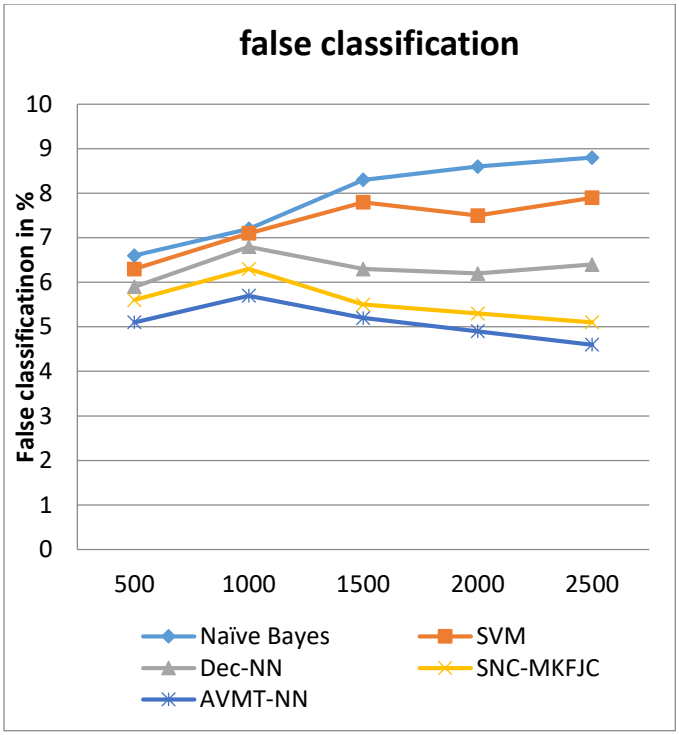

Figure 7 Impact of false classification

Figure 7 shows the comparison of the false classification rate with other methods. The Naïve Bayes method produce 6.6 % , SVM produce 6.3 %, Dec-NN produce 5.9 %, SNC-MKFJC produce 5.6 % and the proposed AVMT-NN produce 5.1 % low level compared to other methods.

Table 8 Impact of false classification

Methods/ Datasets	Comparison of False Classification in %				
	Naïve Bayes	SVM	Dec-NN	SNC-MKFJC	AVMT-NN
500	6.6	6.3	5.9	5.6	5.1
1000	7.2	7.1	6.8	6.3	5.7
1500	8.3	7.8	6.3	5.5	5.2
2000	8.6	7.5	6.2	5.3	4.9
2500	8.8	7.9	6.4	5.1	4.6

Table 8 shows the false rate estimated values under the prediction of classifaction accuracy in type1 dataset. The resultant proves the best performance pointed to matcase with other dissimilar methods.

Time complexity (Tc)

$$=\sum_{k=0}^{k=n} \times \frac{\text{Total Features Handeled to Process in Dataset}}{\text{Time Taken(Ts)}}$$

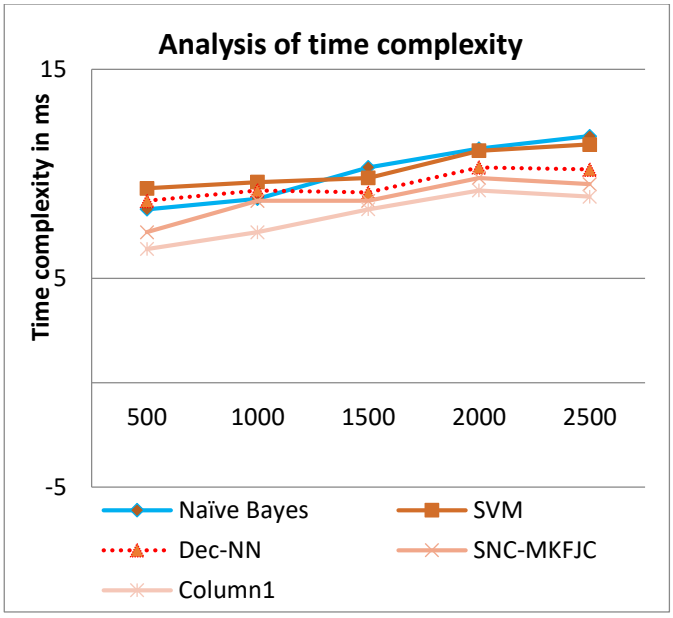

Figure 8 Impact of time complexity

The contrast of total time processing is one of the important parameters for quick analysis for predicting results. Time complexity estimates the impact of execution timing during the process of evaluation. Figure 8 shows the time complexity and its comparisonof differentmethods, proposedAVMT-NNsystem proves the less time taken to complete the process compared to the other systems.

Table 9 Impact of time complexity

Methods/ Datasets	Impact of Time Complexity in Milliseconds (ms)				
	Naïve Bayes	SVM	Dec-NN	SNC-MKFJC	AVMT-NN
500	8.3	9.3	8.7	7.2	6.4
1000	8.8	9.6	9.2	8.7	7.2
1500	10.3	9.8	9.1	8.7	8.3
2000	11.2	11.1	10.3	9.8	9.2
2500	11.8	11.4	10.2	9.5	8.9

The execution time posses the number of records taken by the feature selection and classification process. Table 9 shows the execution of time complexity by various approaches. The proposed method produce high performance in less time consuming compared to the other system. The estimation carried out by asymptomatic notation O(n) in total number of records processed at average mean time. This execution holds the maximum upper limit g(n) and lower limit f(n) at the rate of mean time process. The resultant proves the time complexity 8.9 (ms) performance as well than other methods.

5. Conclusion

To conclude, this proposed diabetic prediction implementation supported has good accuracy in predicting premature conditions. To get a consistent data set, first, the features create relational attribute form of predictions and the features have a negative impact on our neural network for classification. They have decided to choose one of the most useful firefly subset feature selection models to produce the

best selective case for classification accuracy features. Proposed Angular Vector Matrix Transformation Neural Classification has produced higher classification to select the feature with optimal firefly algorithm because the neural network parameters are minimal to improve the classification speed and achieve better results in terms ofsensitivity 95.7 %, specificity 95.6 %, and the classification accuracy 96.2 %. Features of the neural network have produced better decisions on the problem of diabetes classification.

Reference

1. Boiroux, D., Bátora, V., Hagdrup, M., et al.: 'Adaptive model predictive control for a dual-hormone artificial pancreas', J. Process Control, 2018, 68, pp. 105–117
2. Colmegna, P.H., Sánchez-Peña, R.S., Gondhalekar, R., et al.: 'Switched LPV glucose control in type 1 diabetes', IEEE Trans. Biomed. Eng., 2016, 63, (6), pp. 1192–1200
3. Khodakaramzadeh, S., Batmani, Y., Meskin, N.: 'Automatic blood glucose control for type 1 diabetes: A trade-off between postprandial hyperglycemia and hypoglycemia', Biomed. Signal Proc. Control, 2019, 54, p. 101603
4. Goede, S.L., de Galan, B.E., Leow, and M.K.S.: 'Personalized glucose-insulin model based on signal analysis', J. Theor. Biol., 2017, 419, pp. 333–342
5. Batmani, Y., &Khodakaramzadeh, S. (2019). Blood Glucose Concentration Control for Type 1 diabetic patients: A Multiple-model Strategy. IET Systems Biology. doi:10.1049/iet-syb.2018.5049
6. Sierra-Sosa, D., Garcia-Zapirain, B., Castillo, C., Oleagordia, I., Nuno-Solinis, R., Urtaran-Laresgoiti, M., &Elmaghraby, A. (2019). Scalable Healthcare Assessment for Diabetic Patients using Deep Learning on Multiple GPUs. IEEE Transactions on Industrial Informatics, 1–1. doi:10.1109/tii.2019.2919168
7. M. L., ALVA, et al. "The impact of diabetes-related complications on healthcare costs: new results from the UKPDS (UKPDS 84)," Diabetic Medicine, vol. 32, pp. 459-466, 2015
8. Edurne, ALONSO-MORÁN, et al. "Prevalence and quality of care indicators of type 2 diabetes in the population of the Basque Country (Spain)," Avances en Diabetología, vol. 31, pp. 72-79, 2015
9. Guo-li DU, et al. "Metabolic Risk Factors of Type 2 Diabetes Mellitus and Correlated Glycemic Control/Complications: A Cross-Sectional Study between Rural and Urban Uygur Residents in Xinjiang Uygur Autonomous Region," PloS one, vol. 11, 2016.
10. A, A. B., & S, V. (2019). Enhanced optimal insulin regulation in post-operative diabetic patients: an adaptive cascade control compensation based approach with diabetic and hypertension. IEEE Access, 1–1. doi:10.1109/access.2019.2927248
11. Kumarage, P. M., Yogarajah, B., &Ratnarajah, N. (2019). Efficient Feature Selection for Prediction of Diabetic Using LASSO. 2019 19th International Conference on Advances in ICT for Emerging Regions (ICTer). doi:10.1109/icter48817.2019.9023720
12. M. Kumari, R. Vohra, and A. Arora, "Prediction of Diabetes Using Bayesian Network", International Journal of Computer Science and Information Technologies, 2014.
13. Kavakiotis, O. Tsave, A. Salifoglou, N. Maglaveras, I. Vlahavas, I. Chouvarda "Machine Learning and Data Mining Methods in Diabetes Research", Computational and Structural Biotechnology Journal, vol. 15, 2017.
14. Lomte, R., Dagale, S., Bhosale, S., &Ghodake, S. (2018). Survey of Different Feature Selection Algorithms for Diabetes Mellitus Prediction. 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA). doi:10.1109/iccubea.2018.8697452
15. Ke Yan, & Zhang, D. (2014). Blood glucose prediction by breath analysis system with feature selection and model fusion. 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society.
16. C. Turner, "Potential of breath and skin analysis for monitoring blood glucose concentration in diabetes," Expert Rev. Mol. Diagn., vol. 11, no. 5, pp. 497-503, June 2011.
17. S. K. Vashist, "Non-invasive glucose monitoring technology in diabetes management: A review," Anal. Chim. Acta, vol. 750, pp. 16-27, Oct. 2012
18. D. Guo, D. Zhang, L. Zhang, and G. Lu, "Non-invasive blood glucose monitoring for diabetics by means of breath signal analysis," Sens. Actuators: B. Chem., vol. 173, no. 6, pp. 106-113, Oct. 2012
19. K. Yan, D. Zhang, D. Wu, H. Wei, and G. Lu, "Design of a Breath Analysis System for Diabetes Screening and Blood Glucose Level Prediction," IEEE Trans. Biomed. Eng., to be published
20. Thulasi, K. S., Ninu, E. S., & Shiva, K. K. M. (2017). Classification of diabetic patients features using Naïve Bayes classifier. 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT). doi:10.1109/rteict.2017.8256787
21. Imam, M. H., Karmakar, C. K., Jelinek, H. F., Palaniswami, M., &Khandoker, A. H. (2016). Detecting Subclinical Diabetic Cardiac Autonomic Neuropathy by Analyzing Ventricular Repolarization Dynamics. IEEE Journal of Biomedical and Health Informatics, 20(1), 64–72. doi:10.1109/jbhi.2015.2426206
22. Rubaiat, S. Y., Rahman, M. M., &Hasan, M. K. (2018). Important Feature Selection & Accuracy Comparisons of Different Machine Learning Models for Early Diabetes Detection. 2018 International Conference on Innovation in Engineering and Technology (ICIET).
23. M. Gittens, R. King, C. Gittens, and A. Als, "Post-diagnosis management of diabetes through a mobile health consultation application," in eHealth Networking, Applications and Services (Healthcom), 2014 IEEE 16th International Conference on. IEEE, 2014, pp. 152–157
24. D. K. Choubey and S. Paul, "Gamlpnn: a hybrid intelligent system for diabetes disease diagnosis," International Journal of Intelligent Systems and Applications, vol. 8, no. 1, p. 49, 2016.
25. D. K. Choubey, S. Paul, S. Kumar, and S. Kumar, "Classification of pima indian diabetes dataset using naive bayes with genetic algorithm as an attribute selection," in Communication and Computing Systems: Proceedings of the International Conference on Communication and Computing System (ICCCS 2016), 2017, pp. 451–455.