

Article

A Combined Method for Diabetes Mellitus Diagnosis Using Deep Learning, Singular Value Decomposition, and Self-Organizing Map Approaches

Mehrbakhsh Nilashi ^{1,2,*}, Rabab Ali Abumalloh ³, Sultan Alyami ⁴, Abdullah Alghamdi ⁵ and Mesfer Alrizq ⁵

¹ UCSI Graduate Business School, UCSI University, No. 1 Jalan Menara Gading, UCSI Heights, Cheras, Kuala Lumpur 56000, Malaysia

² Centre for Global Sustainability Studies (CGSS), Universiti Sains Malaysia (USM), George Town 11800, Malaysia

³ Department of Computer Science and Engineering, Qatar University, Doha 2713, Qatar

⁴ Computer Science Department, College of Computer Science and Information Systems, Najran University, Najran 55461, Saudi Arabia

⁵ Information Systems Department, College of Computer Science and Information Systems, Najran University, Najran 55461, Saudi Arabia

* Correspondence: nilashidotnet@hotmail.com

Abstract: Diabetes in humans is a rapidly expanding chronic disease and a major crisis in modern societies. The classification of diabetics is a challenging and important procedure that allows the interpretation of diabetic data and diagnosis. Missing values in datasets can impact the prediction accuracy of the methods for the diagnosis. Due to this, a variety of machine learning techniques has been studied in the past. This research has developed a new method using machine learning techniques for diabetes risk prediction. The method was developed through the use of clustering and prediction learning techniques. The method uses Singular Value Decomposition for missing value predictions, a Self-Organizing Map for clustering the data, STEPDISC for feature selection, and an ensemble of Deep Belief Network classifiers for diabetes mellitus prediction. The performance of the proposed method is compared with the previous prediction methods developed by machine learning techniques. The results reveal that the deployed method can accurately predict diabetes mellitus for a set of real-world datasets.

Keywords: Singular Value Decomposition; diabetes; Self-Organizing Map; diagnosis; accuracy; deep learning



Citation: Nilashi, M.; Abumalloh, R.A.; Alyami, S.; Alghamdi, A.; Alrizq, M. A Combined Method for Diabetes Mellitus Diagnosis Using Deep Learning, Singular Value Decomposition, and Self-Organizing Map Approaches. *Diagnostics* **2023**, *13*, 1821. <https://doi.org/10.3390/diagnostics13101821>

Academic Editors: Sameer Antani and Dechang Chen

Received: 9 February 2023

Revised: 10 March 2023

Accepted: 12 April 2023

Published: 22 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Diabetes, also known as diabetes mellitus, in humans is a rapidly expanding chronic disease [1,2] and has major impacts on modern societies [3]. As a person with diabetes mellitus cannot process food properly, glucose builds up in the bloodstream [4,5]. With diabetes, the body cannot produce sufficient insulin (type 1 diabetes) or cannot effectively utilize the hormone (type 2 diabetes) [6–8]. The body stops insulin in type 1 diabetes. In type 2 diabetes, the body's cells do not respond effectively to insulin. This is due to the body's immune system attacking and destroying a part of the pancreas mistakenly. Young people are usually affected by type 1 diabetes, mostly under 30 years of age. Type 2 diabetes, in contrast, often impacts middle-aged and older-aged people and is not fully curable. Some risk factors for type 2 diabetes are family history, age, ethnicity, being obese, having a diagnosis of gestational diabetes, high blood pressure, and obstructive sleep apnea.

The use of machine learning in disease diagnosis has been effective [9–20]. Many algorithms have been developed using supervised and unsupervised learning techniques

and the deployment of real-world datasets for disease prediction [21,22]. Different machine learning techniques have been investigated for diabetes diagnosis in previous studies [23,24]. Although machine learning techniques have presented a powerful performance in knowledge extraction from diabetes medical datasets, the datasets are usually correlation-structured, non-normal, nonlinear, and complex. Therefore, the classification of diabetes disease is a challenging task. Furthermore, the missing values in datasets can influence the prediction accuracy of the deployed method. This has been a critical issue in the datasets for diabetes risk prediction. This study accordingly aims to develop a new combined method for diabetes mellitus diagnosis using deep learning, dimensionality reduction, and clustering approaches. The method is developed using a Self-Organizing Map (SOM), Singular Value Decomposition (SVD), and Deep Belief Network (DBN). We use SOM for data clustering and SVD for dimensionality reduction and missing value prediction. SOM is a competitive, cooperative, and well-known technique for describing the inherent connection between input and output [25]. To find the most important features, a feature selection method is applied to each cluster of SOM. SVD is a robust linear transformation for reducing dimensionality. It enables the space to be arranged to represent the main associative characteristics of the information while ignoring the smaller, less significant influences. The SVD transformation also produces zero mean and intercorrelated features [26]. DBN is used for diabetes mellitus prediction from a set of real-world diabetes mellitus data. DBN is a class of artificial intelligence (AI) computational techniques that enable an algorithm to program itself by learning from a large set of examples, eliminating the necessity to clearly define rules manually. It is powered by developments in computation and large amounts of data [27].

This research is organized as follows. In Section 2, we present related works. In Section 3, the methodology is provided. In Section 4, we introduce the dataset. In Section 5, we perform the data analysis. In Section 6, the discussion of the results is presented. In Section 7, we conclude this work. A list of abbreviations utilized in this research is presented in Table A1 of Appendix A.

2. Literature Review

In medical diagnosis research, disease classification has played a significant part in the medication development of various diseases [28]. Several studies have been conducted on the Diabetes Disease (DD) data classification by utilizing the PID dataset [19,29–31]. By deploying the dataset from the University of California, Irvine (UCI), ML warehouse, scholars have developed multiple approaches for overcoming DD diagnosis issues and enhancing the accuracy values. In DD classification research, various classification models and taxonomies have been presented in recent years, leading to a complex mix of detailed and imprecise terminologies. In the following, we will present a summary of previous studies that are present in this field.

Early diagnosis of retinal abnormalities has the chance to avoid diabetes, in addition to being vital for avoiding the development of diabetic retinopathy and/or diabetic macular edema (DME), as well as other abnormalities that could result in visual impairment. Existing DR screening techniques rely on retinal fundus imaging and conventional evaluation. Even so, such strategies are exceedingly laborious and expensive, as highly qualified experts are required for evaluation purposes [32].

In the study presented by Polat et al. [33], the authors designed a new approach for DD detection based on the GDA and the LS-SVM techniques. A novel cascade learning system was presented based on the above-mentioned techniques. The designed system entailed two stages: (1) GDA, which was used to preprocess the initial data and discriminate between normal and diabetes cases, and (2) LS-SVM, which was utilized in the second stage to categorize the diabetes datasets. The outcomes presented a classification accuracy of 82.05%, which was considered an encouraging outcome when compared with prior findings reported by other classification techniques.

A study by Aslam et al. [34] used GP for diabetes classification. In this study, new features were produced using combinations of the current features of the disease, without proceeding knowledge of the probability distribution. The designed method had multiple basic steps. First of all, the features were selected and utilized to generate an ordered set of initial features. By referring to the ordered set, features were added one after another to constitute several subsets of the initial features. Following that, new features were produced from each subgroup of initial features using GP-CPS to utilize the strengths and weaknesses of GP-generated features. In the final step, the performance of the produced features was evaluated using the K-NN and SVM techniques. The outcomes of the study compared to other approaches presented a superior performance. A new approach was introduced to classify medical records by Kahramanli and Allahverdi [35]. In addition, a hybrid NN system, which entailed the ANN method and FNN approach, was designed. Two real-time datasets were utilized for checking the suitability of the proposed approach. In order to assess the performance of the suggested approach, several metrics were used. The proposed approach achieved ACC values of 84.24% and 86.8% for the two datasets. The proposed method presented the best outcome compared with proceeding outcomes obtained from similar previous studies.

A new method was presented by Erkeymaz and Ozer [36] for DD diagnosis based on the SW-FFANN approach, which was the first study that deployed this approach for DD. The authors developed a small-world network by following the Watts–Strogatz method, utilizing the proposed structure for DD classification, and contrasting the research outcomes with regular or conventional FFANN. The authors indicated that SW-FFANN presented better outcomes than the conventional FFANN in terms of output correlation, output error parameters, and classifier performance. When compared with other approaches that utilized the Pima Indians Diabetes database, the deployed approach presented a better performance.

Fuzzy methods have become one of the prominent solutions for classification issues, as they can enhance the interpretability of the outcomes and present more insight into the classifier scheme and decision-making procedure. A study by Ganji and Abadeh [37] used a system to extract a group of fuzzy conditions for DD classification based on the ACO technique. The new system had novel features that distinguish it from other similar systems. The classification accuracy was 84.24%, which indicated that the proposed system outperformed multiple well-known systems. Another study that deployed fuzzy methods was presented by Dogantekin et al. [38]. An intelligent system for DD diagnosis was developed based on the LDA and ANFIS techniques by Dogantekin et al. [38]. The architecture of the DD diagnosis system contained two main components: LDA, which was utilized to extract feature variables between healthy and DD patient data. Four main metrics were used in the evaluation of the system: SE and SP, ACC, and CM. The results indicated that the classification accuracy of the system was around 84.61%. A comparative Pima-DD diagnosis was accomplished by Temurtas et al. [39]. For this purpose, a MLNN architecture, which was trained by the LM algorithm, and a PNN architecture were utilized. The outcomes of the study were assessed compared to the outcomes of prior reports concentrating on the DD diagnosis and utilizing the same ML database. The classification accuracy of MLNN with LM obtained by the study using correct training was better than those obtained by other studies for the conventional validation method. The outcomes, based on PID, that resulted from using PNN architecture, presented a good performance compared with the outcomes from other research, especially for similar approaches. However, the study presented far less classification accuracy ($10\times$ FC) (66.78%) compared to the study by [40]. Another study by Hayashi and Yukita [29] utilized the PID dataset by incorporating a rule extraction algorithm and sampling selection techniques to achieve high-quality results in terms of ACC and conciseness. The ACC value was 83.83% by using the rule extraction algorithm, while the sampling selection technique presented better results in terms of the ACC value and conciseness. The novelty of this study is reflected by its ability to overcome the black box model by adopting the rule extraction method to present such explanations.

SVM was used by many DD diagnosis researchers [19,41]. In a study by Çalişir and Doğantekin [41], a DD diagnosis system based on LDA and MWSVM was presented. The architecture of the proposed system entailed three steps: LDA was used in the feature extraction and feature reduction stage; MWSVM was used in the classification step; and the third step entailed the evaluation process of the result based on the following metrics: SE, SP, ACC, and CM. The ACC of the proposed method was obtained at about 89.74%. Compared to previous studies that used the same approach, it presented a better performance, such as [33] presented an ACC of 78.21% and [33] presented an ACC of 79.16%.

Another study by Nilashi et al. [19] developed a DD diagnosis system based on ML techniques, particularly the EM, PCA, and SVM approaches. EM was used basically for clustering, PCA was used for removing the noise, and SVM was used for performing the classification. The incremental PCA, along with the incremental SVM, were used for the incremental learning of data. The experimental findings on the PID dataset presented an enhancement of the ACC value compared to the conventional methods. The study stressed the importance of the deployment of incremental learning on additional datasets to enhance the computation time. Aiming to improve the DD diagnosis accuracy, Sisodia and Sisodia [30] used three ML techniques: DT, NB, and SVM. Experiments were conducted on the PID dataset. The assessment of the research outcomes was conducted, focusing on several metrics of PR, ACC, F-Measure, and REC. The ACC value was measured based on correctly and incorrectly classified instances. The findings indicated that, among the inspected algorithms, NB had the highest accuracy value of 76.30%. Still, the study is limited in terms of the assessment of the presented results compared to other related works. NN approaches have been utilized in many studies regarding DD diagnosis [31,35,36]. ML approaches were utilized in a study presented by Nilashi et al. [31] by designing a DD classification system that combined the SOM, PCA, and NN techniques. SOM was used for clustering, PCA was used for removing the noise, and NN was used for the DD diagnosis. The experimental outcomes on the PID dataset presented an enhancement of the accuracy over previous reports in the same context. The deployed approach helped in enhancing the accuracy significantly compared to other approaches, such as general regression neural networks by [42], GDA-LS-SVM by [33], and SW-FFANN by [36].

Another study that adopted the NN approach was presented by Edla and Cheruku [43]. The study adopted RBFNN and utilized a Bat-based clustering algorithm. The designed approach was evaluated on the PID dataset and presented outcomes that surpass other variants of RBFNN in several studies, such as the Dunn index-RBFNN by [44] and Bee-RBF model by [45]. Another study presented by Khan et al. [46] aimed to classify the DD medical status among patients using ANN approaches, particularly the LM algorithm on the PID dataset. The result of the study was assessed based on measures of SE and SP for various learning rules. The study indicated that the LM algorithm surpassed other approaches in terms of the considered evaluation measures. It also helped in addressing the convergence issue in some other algorithms. In a study by Maniruzzaman et al. [47], the GC approach was used to overcome the non-normality, inherent correlation structure, and nonlinearity issues. The outcomes were assessed compared to the current methods, such as NB and LDA, according to several metrics of ACC, SE, SP, PPV, NPV, and ROC, and presented encouraging outcomes compared to other techniques that were deployed in the literature, such as ANN and AIS by [48] and SVM by [49]. DL techniques were used by many studies in the context of DD classification [28,50]. In a study by Kannadasan et al. [28], a framework based on DL was designed for DD classification. Several metrics were used for the assessment of the designed system, and an overall ACC value of 86.26% resulted. In a work presented by Alhassan et al. [50], a huge dataset for DD was presented (KAIMRCD). The dataset entailed comprehensive information about 14,000 cases. Aiming to achieve an accurate T2DM diagnosis, the dataset was processed as a time series and then explored through a TP-DL approach. Additionally, LSTM and GRU were trained on the dataset and presented a high ACC value of over 97%.

PK-SVM was used by Longato et al. [51] to differentiate between individuals impacted by IGT vs. T2D based on CGM-mined GV instances and focus on several essential variables: age, sex, and BMI. The outcomes presented an acceptable ACC value of 87.1%. In a study by Rahman et al. [24], a novel DD classifying model based on Conv-LSTM was developed. In addition, three well-known approaches: CNN, T-LSTM, and CNN-LSTM were assessed, and the outcomes were compared against the outcome of the proposed model using the PID dataset. The evaluation findings revealed that Conv-LSTM presented the best performance among the assessed approaches, with the highest ACC value of 97.26%. Devi et al. [52] presented a combined method that utilized the FF clustering technique and SMO technique for classifying DM. To achieve the research goal, the FF algorithm was used to segment the data into multiple segments. As the dataset size reduced, the result of the time used for the computation was minimized noticeably. The output of the FF technique was presented to the SVM, which categorized the cases into diabetic and non-diabetic, based on the PID and achieved a high ACC value (99.4%). In a study by Bhatia et al. [23], an effective home-based supervision system for DD classification was presented. The system entailed four basic layers: DDA, DDC, DME, and DPDM layers to classify, extract, predict, and classify DD. In addition, RNN was used for analyzing the prediction regarding the probabilistic quantification of UbD supervision considering the LoDI and DIM. Additionally, UbD existence was viewed based on the SOM process. The research outcomes were assessed based on several measures and presented enhanced values compared to other studies.

In summary, assessing the efficacy of DD classification is a complicated issue that should be tested on newer and more comprehensive diabetes datasets in future research, so that the most highly accurate rules can be extracted for diagnosis [29].

3. Methods

The method of this study is developed for the diagnosis of diabetes mellitus. The general framework of the proposed method is presented in Figure 1. The method uses SVD for missing value prediction, SOM for clustering the data, STEPDISC for feature selection, and an ensemble of DBN classifiers for diabetes mellitus prediction. These techniques are introduced in the following sections. In the first step of the proposed methodology, the data is preprocessed. Then, the data clustering is performed. The clustering of SOM is performed with the aid of PCA, as the noise of the data may impact the classification accuracy. The aim of performing clustering is to provide more accurate classification results. In addition, the prediction model of segments that include diabetes mellitus is more effective than the model that considers the whole dataset for diabetes classification. In addition, in the proposed approach, we incorporate the feature selection strategy to find the most important features to predict diabetes mellitus by the use of ensembles of DBN classifiers.

3.1. SVD

In this study, SVD as a dimensionality reduction technique was used for the missing value prediction. The procedure for missing value prediction by SVD is presented in Algorithm 1. This method can predict the missing value in the datasets through five main steps. To do so, we first convert the data in each cluster to the dense matrix $\mathbf{B}^{m,n}$. Then, we perform the normalization procedure $\mathbf{B}^{m,n}$. In the third step, the SVD technique is applied to the matrix obtained in the second step. The obtained matrix Z is used for the approximation of Z , Z_d . SVD finally predicts the missing value in the last step.

Algorithm 1: Missing value prediction by SVD.

- Converting data to the dense matrix $B^{m,n}$.
- Normalization using $Z_{ij} = \frac{B_{ij} - \bar{B}^j}{\sigma_j} \bar{B}^j = \frac{1}{m} \sum_{i=1}^m B_{ij}, \sigma_j^2 = \frac{1}{m-1} \sum_{i=1}^m (B_{ij} - \bar{B}^j)^2$ where \bar{B}^j is the average value and σ_j indicate the SD for B^j .
- SVD technique on Z is applied.
- Z_d is computed for approximation of Z.
- Missing value is calculated using $\bar{B}^j + \sigma_j(Z_d)_{ij}$.

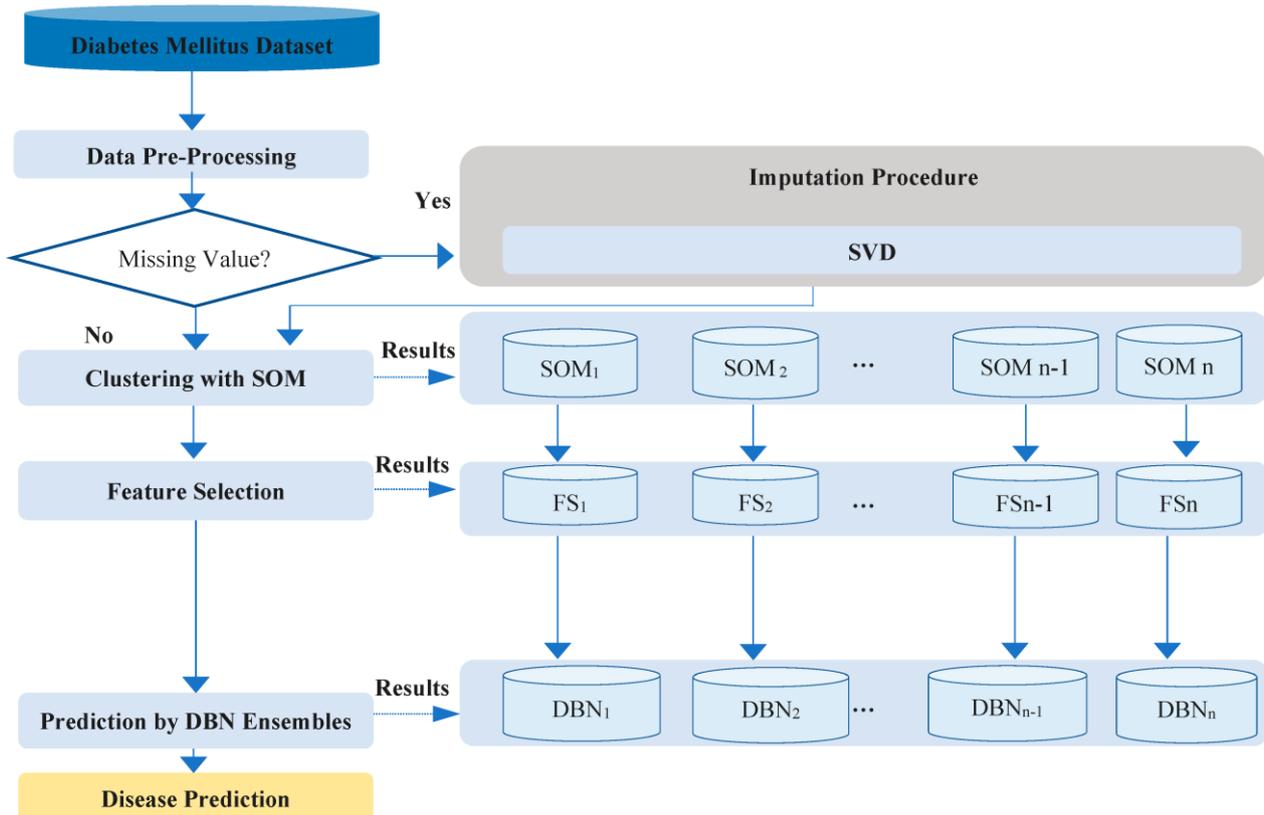


Figure 1. Research method.

3.2. SOM

Kohonen Self-Organizing Maps (SOM) [53] have been utilized as a robust technique for data clustering. The SOM structure includes one input and one output layer [54]. The objective of SOM is to transfer all input data objects with n dimensions to the output in a manner in which the objects are related to each other [55,56]. The SOM is able to perform unsupervised training for the datasets, in which the target is clustering the data [14,15,57–61]. The Euclidean distance [62,63] from the input vector to all nodes is computed when the input vector is given to the network. All weight vectors of the SOM network are randomly initialized, and from the training dataset, an input pattern X is chosen. Through the comparison of the Euclidean distances, the similarity to the input vector is found, and this will enable the network to identify the winning node, the Best Matching Unit (BMU), which is closest to the input vector. Three processes, which are competition, cooperation, and adaptation, are essential in the SOM algorithm. In the competition step, the winning neurons for observation are identified. In the cooperation process, SOM determines a topological neighborhood, and cooperation is ensured between nodes. In the adaptation process, according to the example, the synaptic weights are arranged for the neurons affected by competition.

3.3. DBN Technique

DBN was designed by Hinton et al. [64] to overcome the vanishing gradient issue presented in prior research [65]. DBN is a robust multilayered generative design that is constructed from pretrained layers, and it is categorized as a DNN approach [17]. DBN entails multiple RBMs that are composed of a Visible Layer (VL) and a Hidden Layer (HL); the first is considered as an input component, while the second is regarded as an output part. Fully linked nodes in different layers are presented without links among the nodes in the internal layers. RBM attempts to learn the distribution of the probability from VL to HL using an Energy Function (EF) [17]. DBN entails the pretraining step (unsupervised) of the deep RBMs and the fine-tuning step (supervised) of the classification layer. Following the prior training step, the standard backpropagation formula is utilized in a controlled style to refine all the weights to achieve a specific role. Weights initialization of the network using DBN is applied effectively in several domains related to artificial intelligence [66]. DBN presents a good feature extraction result; hence, it is suitable for detecting the features from the data. As DBN is a completely linked structure, it eases the analysis of the data more than any other DNN.

As presented in Figure 2, each RBM is structured from a VL that entails the V-units, $v = \{v_1, v_2, \dots, v_i\}$, and a HL that entails the H-units, $h = \{h_1, h_2, \dots, h_j\}$.

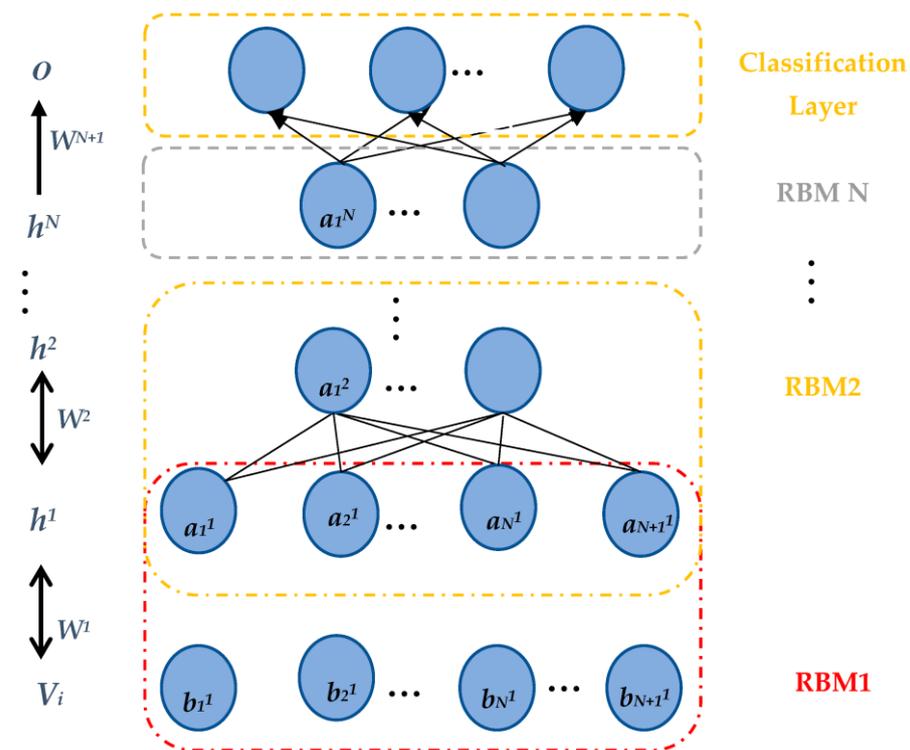


Figure 2. The DBN architecture.

Considering the model variables of DBN, $\theta = [W, b, a]$, the energy function can be presented as:

$$E(v, h; \theta) = -\sum_{i=1}^I \sum_{j=1}^J w_{ij} v_i h_j - \sum_{i=1}^I b_i v_i - \sum_{j=1}^J a_j h_j \tag{1}$$

where w_{ij} is the weight of the link among the visible part v_i , in which the total value is I , and hidden part h , which total value is J , and b_i and a_j indicate the bias requirements of the

visible and hidden parts, respectively. The joint distribution over all the nodes is measured regarding the function of the energy, $E(v, h; \theta)$, as:

$$p(v, h; \theta) = \frac{\exp(-E(v, h; \theta))}{Z} \quad (2)$$

where $Z = \sum_{h,v} \exp(-E(v, h; \theta))$ is the function of the partition. The conditional probabilities of the hidden and visible nodes h and v can be measured as:

$$p(h_i = 1|v; \theta) = \delta \left(\sum_{i=1}^I w_{ij} v_i + a_j \right) \quad (3)$$

$$p(v_i = 1|v; \theta) = \delta \left(\sum_{j=1}^J w_{ij} h_j + b_i \right) \quad (4)$$

whereby δ is a logistic equation, $\delta(x) = 1/1 + \exp(x)$.

The joint probability is optimized by training RBMs. Several RBMs are stacked to construct the DBN, whereby the output of the l th layer (H nodes) is utilized as the input of the $l + 1$ th layer (V nodes). The training procedure is indicated by basic steps: prior training and fine-tuning. In the prior training step, the input data are provided to the VL of the initial RBM and converted into the HL, which is continually processed in the following RBM. Afterward, the layer-to-layer unsupervised training is completed and the DBN learns the features from the given input, then the features are provided to the classifier layer of the DBN. Eventually, fine-tuning is completed on the classification layer to improve the DBN performance.

3.4. Featured Selection by STEPDISC

Based upon several quantitative variables and classification variables, a step-by-step discriminant analysis is carried out by the STEPDISC procedure to select a subset of the quantitative variables to be used in class discrimination [67]. The group of variables that comprise each class is supposed to be normally multivariate, with a common covariance matrix. Observations containing missing values are omitted from the analysis, which may impact the accuracy of the classification models. Forward and backward strategies are used in STEPDISC. In this research, the backward strategy is used. For the stopping rule, the significance level of the F-test from the covariance analysis is used.

4. Dataset

The diabetic dataset includes 768 female patients. It has 500 controls and 268 diabetic patients. This dataset was taken from the UCI (University of California, Irvine) in November 2019. The National Institute of Diabetes and Digestive and Kidney Diseases provided the original data for this dataset. The dataset's goal is to diagnose whether or not a patient has diabetes based on certain diagnostic measurements included in the dataset. The patients are at least 21 years old of Pima Indian heritage. In this dataset, 374 of these patients have zero serum insulin levels, 27 of them have zero body mass indexes, 35 patients have zero diastolic blood pressure, the skinfold thickness is zero for 227 patients, and 5 patients have zero glucose levels. We consider these zero values as missing values. According to the WHO standards, these details can be used to detect the critical symptoms that can be used to predict diabetes in patients. The target variable is a binary variable indicating whether or not a patient has diabetes, with 1 indicating that the patient has diabetes and 0 indicating that the patient does not have diabetes. Demographics of the diabetic patient cohort are shown in Figure 3 and Table 1. This dataset has been widely used in research on diabetes prediction and has been used as a benchmark dataset for developing and evaluating machine learning algorithms for diabetes prediction.

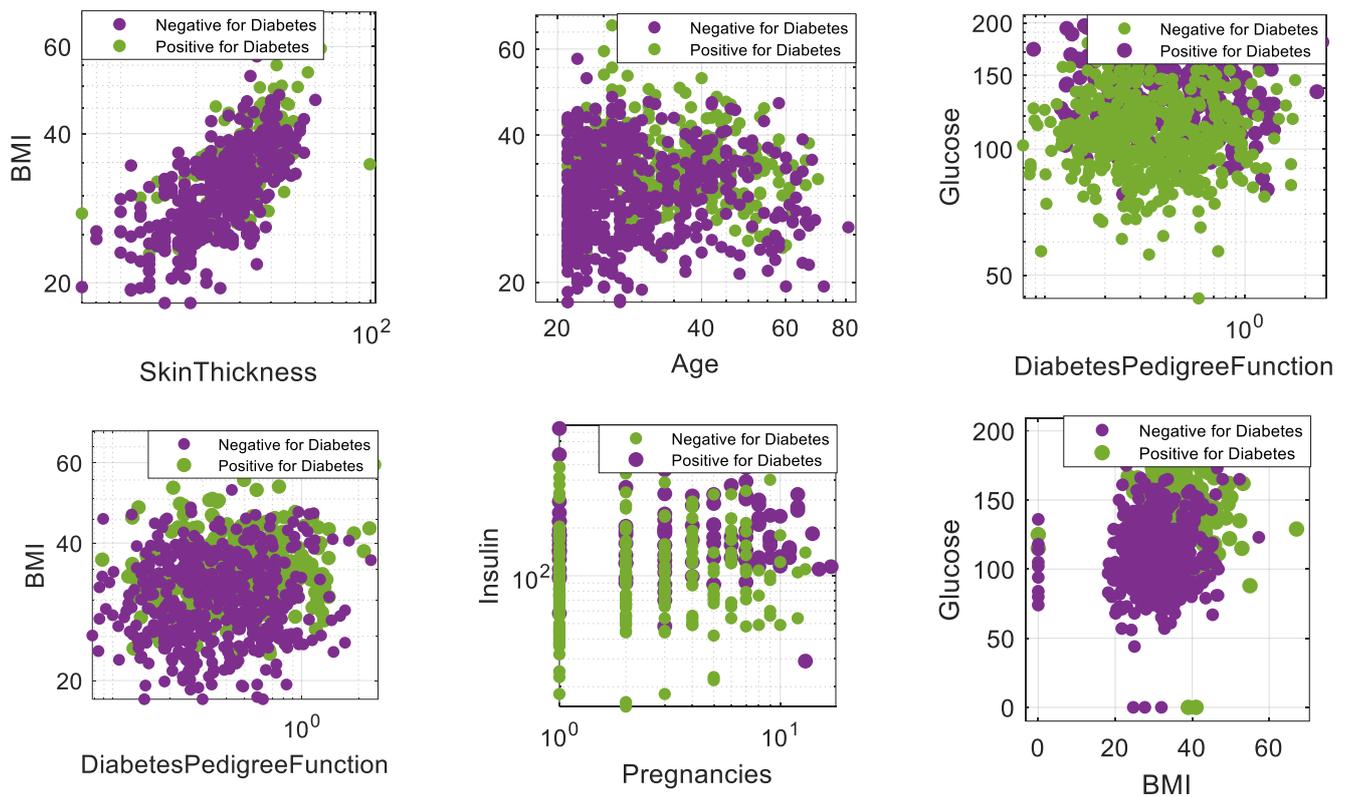


Figure 3. Visualization of the features of the diabetic dataset.

Table 1. Demographics of the diabetic patient cohort.

SN	Attributes	Descriptions	Attributes Type	Mean ± SD
1	Pregnant	Number of times pregnant	Continuous	3.84 ± 3.36
2	Glucose	Plasma glucose (2 h)	Continuous	121.67 ± 30.46
3	Pressure	Diastolic blood pressure (mm Hg)	Continuous	72.38 ± 12.10
4	Triceps	Triceps skin fold thickness (mm)	Continuous	29.08 ± 8.89
5	Insulin	Two hours serum-insulin (μ U/mL)	Continuous	141.76 ± 89.10
6	Mass	Body mass index (weight in kg/(height in m) ²)	Continuous	32.43 ± 6.88
7	Pedigree	Diabetes pedigree function	Continuous	0.47 ± 0.33
8	Age	Age (years)	Continuous	33.24 ± 11.76
9	Class	Diabetic (500) vs. Control (268)	Categorical	-

5. Results and Methods Comparison

In the first step of the data analysis, we use SOM for data clustering. The results for SOM clustering are presented in Figure 4. The results are presented on different PCs of the PCA. Overall, six clusters were discovered by the SOM algorithm. In fact, SOM2×3 resulted in the best map quality for data clustering. The number of samples in Clusters 1–6 is, respectively, 93, 130, 104, 242, 132, and 67. They are visualized by PC4 and PC5, as shown in Figure 5. In this figure, pregnancies, glucose, blood pressure, skin thickness, insulin, BMI, diabetes pedigree function, and age are also visualized in different clusters through different PCs. After data clustering, we apply the SVD algorithm for the missing value prediction. An example of how the SVD algorithm can predict the missing values in the diabetic patient dataset is presented in Figure 5. The procedure is able to perform the prediction through the decomposition of matrices, which include the data of each cluster. The example is provided for low-rank approximation ($D = 2$) of decomposed matrices. The SVD result for diabetes mellitus diagnosis is presented in Figure 6.

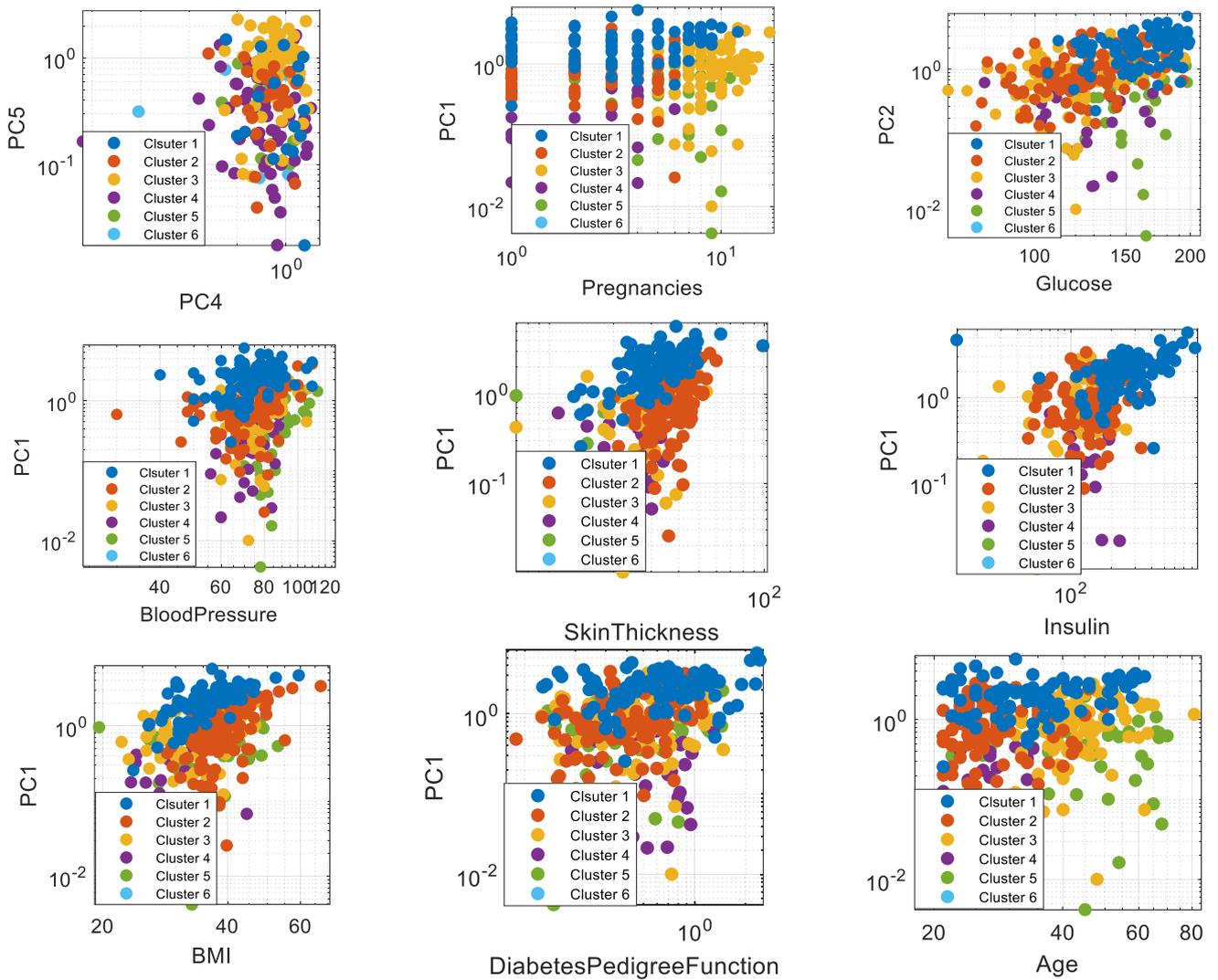


Figure 4. The results of SOM clustering.

In this paper, the classification models are constructed based on the DBN. In each cluster, the bottom of the model is stacked by RBM layers. In addition, the expected output variables are at the top of the model, which is the classification layer. To overcome the data processing complexity problem, one-class-one-network is adopted to improve the ability of the neural network. This approach was suggested by several researchers in different contexts, such as automatic handwriting recognition [68], speech recognition [69], and face recognition [70]. The global one-class-one-network aims to coordinate all subnetwork outputs in classification problems. This can present an appropriate choice for the tradeoff between accuracy and complexity. We applied DBN onto each cluster to construct the classification molds for diabetes mellitus diagnosis. The data for each cluster was divided into training and test sets. To avoid Gradient Diffusion (GD), we applied DBNs’ pretraining and fine-tuning procedures. Selecting an appropriate number of hidden layers and the number of their nodes is important in constructing DBN for the classification task. In this study, the number of hidden layers in DBN is fixed to two. To train the DBN to construct the models, the number of neurons in the input layer was set equally to the number of input variables in the dataset. The number of output layer neurons was set to the number of dataset classes equally. With a batch size of 100, the unsupervised training and fine-tuning algorithms were performed for 200 epochs in each run. In Figure 7, RBMs’ iterative learning process and DBN training process are presented.

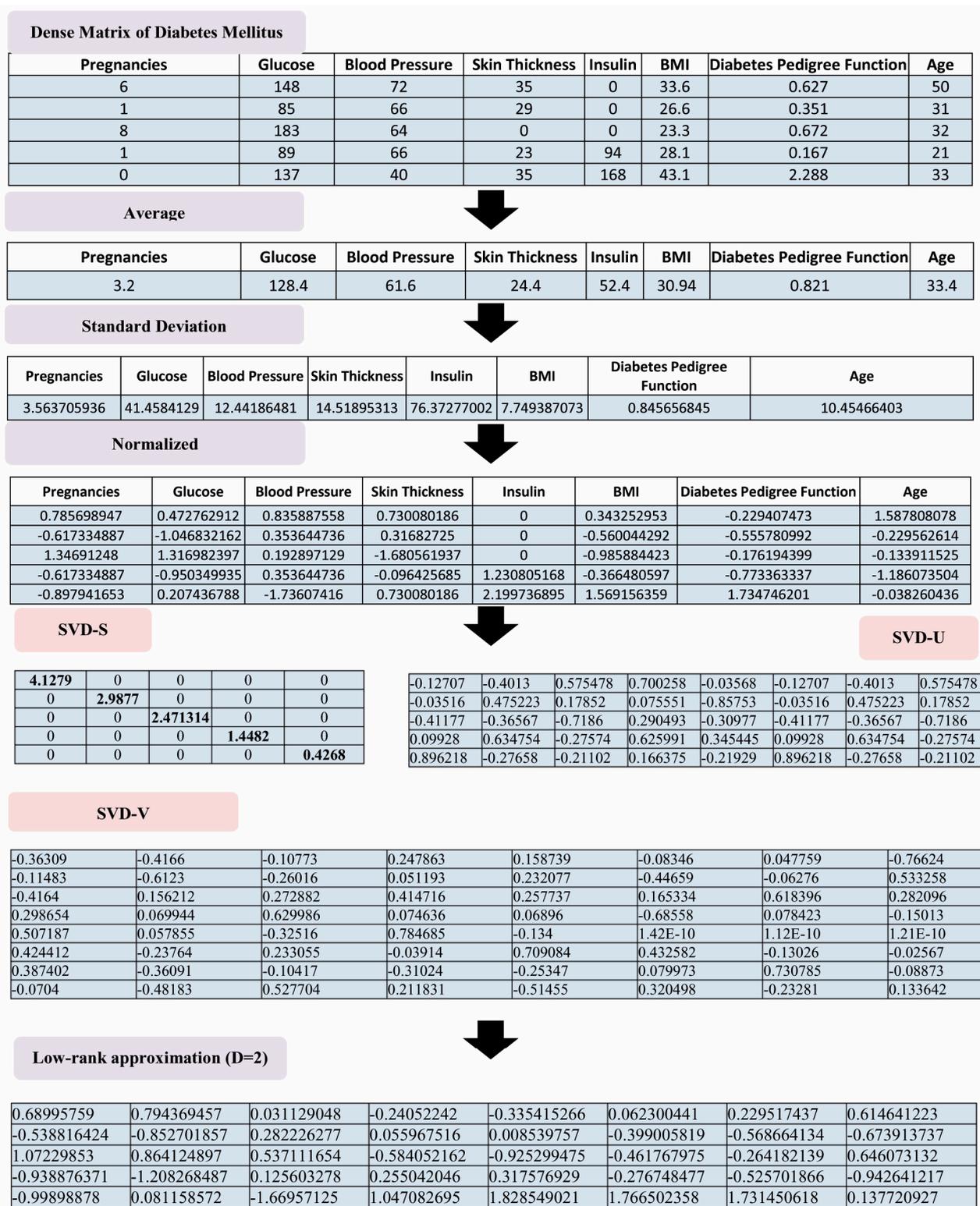


Figure 5. An example of performing the SVD algorithm for missing values prediction.

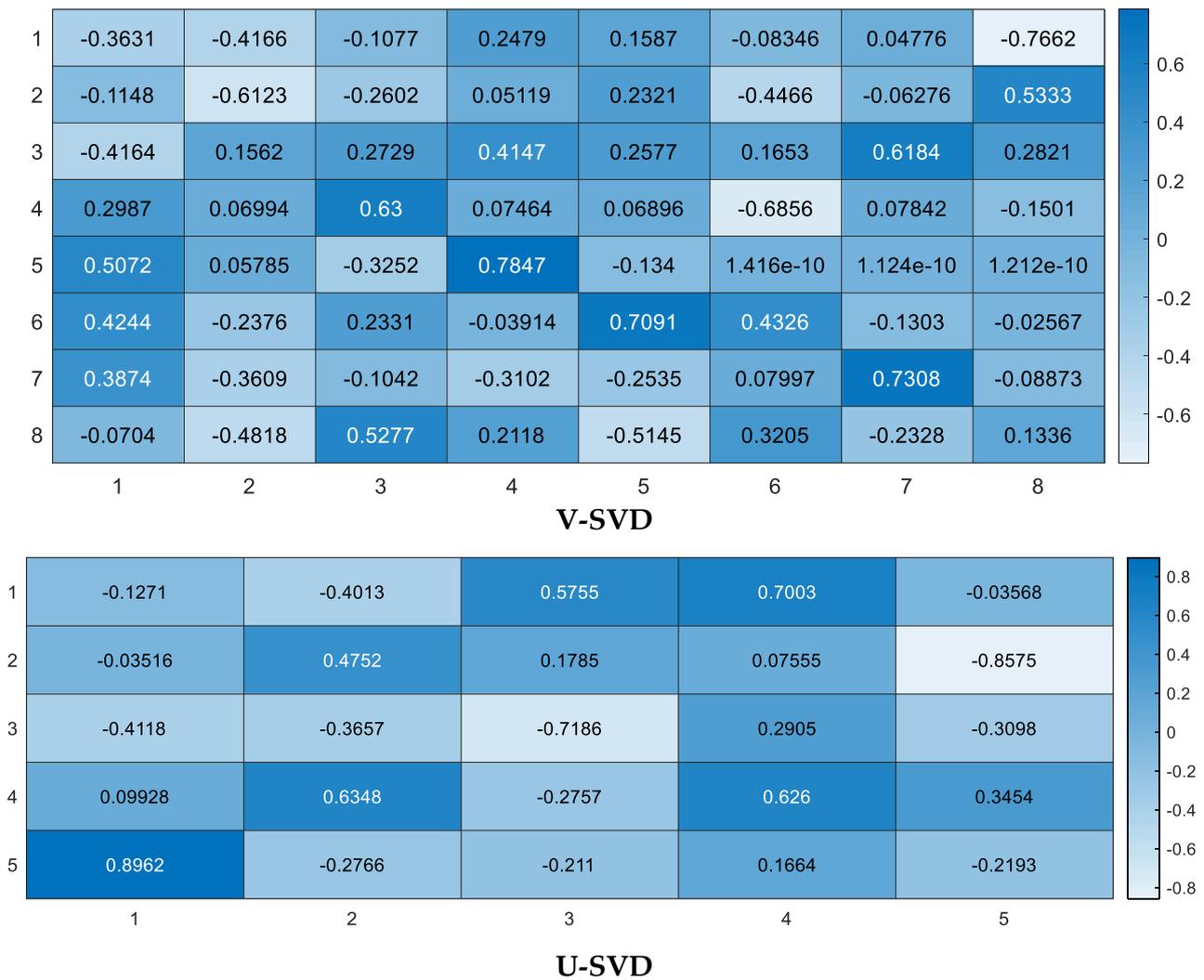


Figure 6. SVD results for diabetes mellitus diagnosis.

The results of our method are evaluated through the accuracy metric. In Figure 8, the formulas to calculate the accuracy are presented. It is obtained through the confusion matrix, which includes True Positive (TP), False Negative (FN), False Positive (FP), and True Negative (TN).

The proposed method is compared with the other classifiers combined with the SOM clustering. The training errors of all the classifiers, DBN, SVM, NN, and LDA, in six clusters are presented in Figure 9. The results are presented for 200 epochs. It is found that the proposed method, SOM + DBN, has better training errors with faster convergence in all clusters.

In addition, it can be seen from the results in Figure 9 that there is no fluctuation in all methods during the learning process, which means that the architecture provided by clustering and classifiers may be stable enough to learn an accurate classification model for the diabetes mellitus diagnosis. We further evaluated the methods by the accuracy metric. The results for all methods are provided in Table 2.

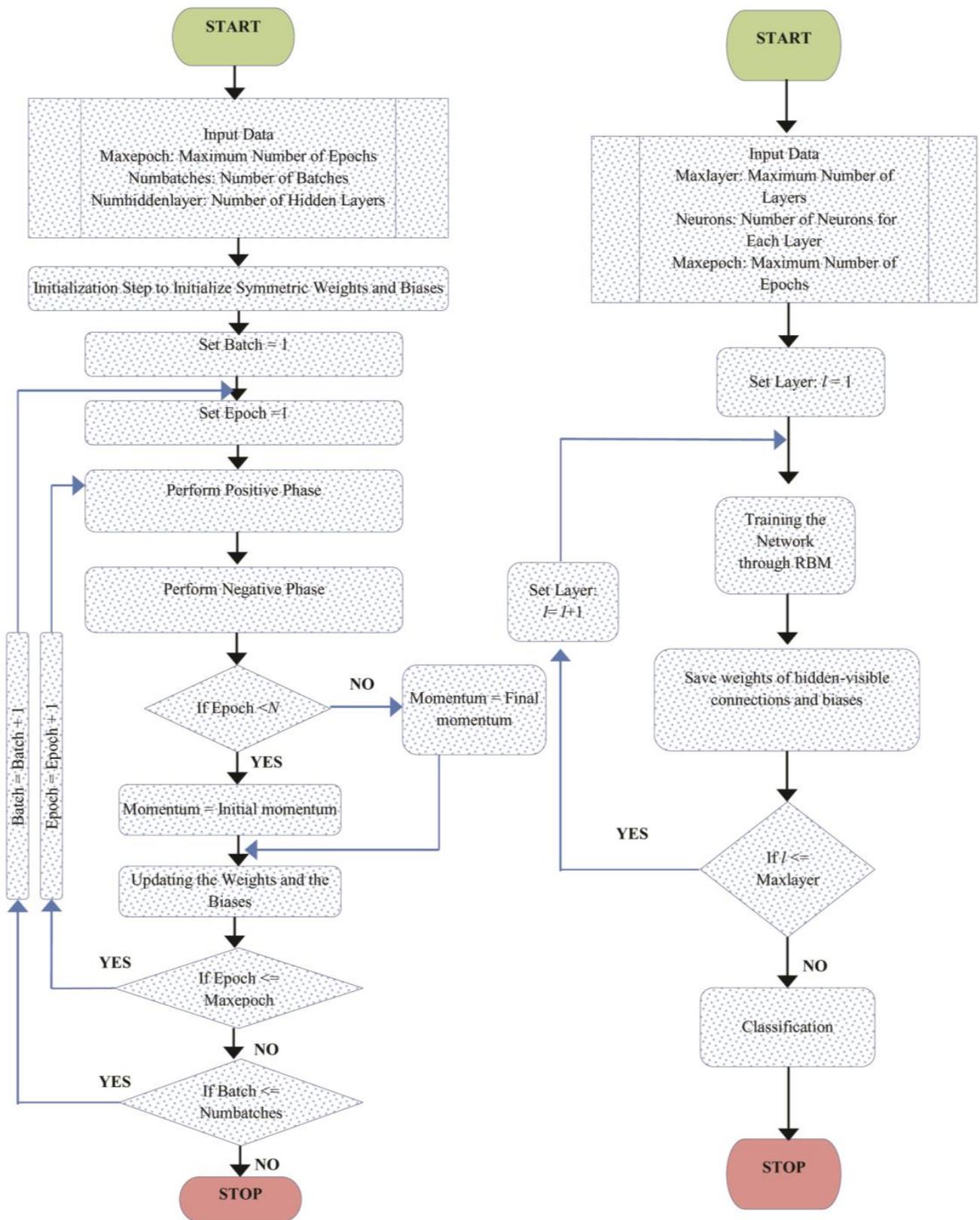


Figure 7. RBM iterative learning process and DBN training process.

		Predicted Class		
		Positive	Negative	
Actual Class	Positive	True Positive (TP)	False Negative (FN)	Sensitivity= $\frac{TP}{TP+FN}$
	Negative	False Positive (FP)	True Negative (TN)	Specificity= $\frac{TN}{TN+FP}$
		Precision= $\frac{TN}{TP+FP}$	Negative Predictive Value= $\frac{TN}{TN+FN}$	Accuracy= $\frac{TP+TN}{TP+TN+FP+FN}$

Figure 8. Confusion matrix.

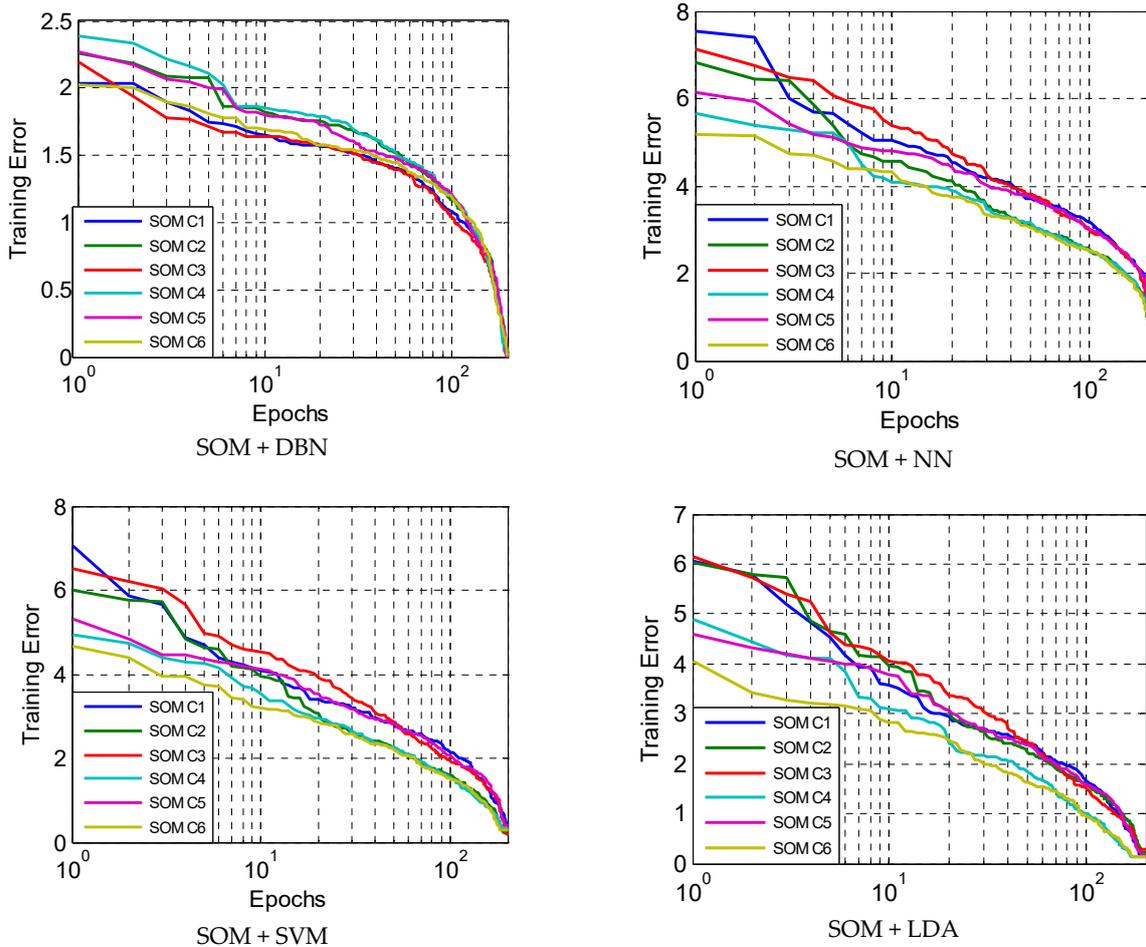


Figure 9. Methods comparison.

Table 2. Comparison of the proposed method with other classifiers for DM (Diagnosis Methods) comparison.

Method	Reference	Accuracy
General Regression Neural Network	[42]	80.21%
GDA-LSSVM	[33]	79.16%
MWSVM	[41]	89.74%
SW-FFANN	[36]	91.66%
SOM + SVM	This study	90.73%
SOM + LDA	This study	91.83%
SOM + NN	This study	84.62%
SOM + DBN Ensembles	This study	98.32%

In addition, our method is compared with the General Regression Neural Network [42], GDA-LSSVM [33], MWSVM [41], and SW-FFANN [36]. The results show that the SOM + DBN ensembles approach provides the best accuracy results (98.32%) for diabetes mellitus classification in relation to the General Regression Neural Network [42] (80.21%), GDA-LSSVM [33] (79.16%), MWSVM [41] (89.74%), and SW-FFANN [36] (91.66%). Furthermore, it is found that the clustering method combined with the classifiers has achieved more accurate classification results.

6. Discussion

Diabetes is a significant disease, because it affects a large number of people and, if not properly managed, can lead to serious health complications [71]. Diabetes is a chronic disease that occurs when the body's ability to regulate blood sugar levels is compromised. Diabetes is classified into two types: type 1 and type 2 [72]. Type 1 diabetes occurs when the body's immune system attacks and destroys insulin-producing cells in the pancreas, whereas Type 2 diabetes occurs when the body becomes insulin-resistant or fails to produce enough insulin to regulate blood sugar levels. Diabetes can cause a variety of health problems, including cardiovascular disease, nerve damage, kidney damage, vision damage, and foot damage. It can also cause other health issues, such as high blood pressure, high cholesterol, and obesity.

Using state-of-the-art computational approaches to aid in complicated medical care choices has been indicated since the 1970s [73]. Still, medical approval and employment of these approaches have moved slower than intended [74]. This can be due to several causes that range from the restricted use of research outcomes to the limited computational performance of these approaches. Hence, with the availability of online medical datasets, it is critical to deploy efficient computational approaches in medical contexts such as disease diagnosis.

Designing new disease diagnosis approaches is a comprehensive process that entails the integration of theory, data gathering, and appropriate data analysis [75]. In the design of the disease diagnosis approach, a group of indicators that reflect the folds of disability resulting from the disease is essential [76]. Diabetes patients need to control and manage their health day by day, which can influence their daily activities. The single method for diabetes patients to tolerate DD is to maintain their blood sugar at normal levels, which can only be managed if the used treatment entails the appropriate diet, taking oral diabetes treatment or some kind of insulin, and practicing suitable exercises [77]. Additionally, the medication for DD is very complex and costly for medical centers [31]. As there are several significant aspects to document regarding the DD that can aid health staff in reaching the appropriate choice about treatment, utilizing the appropriate method for the disease diagnosis is of great importance.

ML addresses the advancement of techniques that enable machines to learn. The obstacle that faces researchers in this regard is to design algorithms that can gather a set of patterns and present new presumptions from the primary data without the need for human intervention. For ML, classification concerns dividing a group of data into multiple classes, referring to the training outcomes of a subgroup of data that have a predefined classification.

Clustering is a procedure that entails dividing a group of observations into multiple reasonable segments based on specific metrics of resemblance among each segment. The clustering approach has been utilized in several disease diagnosis systems [19,78–80], which indicates its wide utility as one of the procedures in experimental medical data analyses.

As a particular fold of ML, DL techniques have been used for exploring various medical datasets to classify diseases and predict their effects [81,82]. DL classification systems have achieved multiple aims in several medical contexts, such as image and unstructured data processing and to aid healthcare staff in their decisions [17]. Many DL systems have accomplished high accuracies at various diagnostic functions among multiple medical conditions [83–85]. The performance of the DL approach highly relies on the optimum measures of variables. The optimum measures of variables are achieved by tuning with various measures and observing the performance where the approach presents a better accuracy. The main aim of this research is to design a new approach for DM diagnosis by combining several techniques and enhancing the disease prediction accuracy. The main contribution of this research is the design of a new approach that incorporates dimensionality reduction, clustering, feature selection, and DBN ensemble techniques to enhance the accuracy prediction of the diabetes risk. The advantage of the current approach can be indicated by the best accuracy outcome compared to other techniques that used the SOM and DBN techniques using the UCI dataset. To summarize, the outcomes of our research on a UCI dataset indicated the efficiency of combining the dimensionality reduction, clustering, and DBN techniques in enhancing the diagnosis accuracy of the diabetes risk.

7. Conclusions

ML has been effectively deployed in the diagnosis of diabetes. This study presented a novel integrated approach using machine learning techniques for early-stage diabetes risk prediction. SVD, SOM, STEPDISC, and DBN ensemble techniques were integrated to develop this method. SVD was used for dimensionality reduction and missing value prediction. SOM was used to cluster the data. STEPDISC was used for feature selection. DBN ensemble was used for diabetes mellitus prediction. To assess the proposed approach, multiple experimental analyses were performed on a real-world dataset taken from UCI. The deployed approach was assessed compared to other approaches in the literature and the methods combined with the SOM clustering. For 200 epochs, the proposed approach, SOM+DBN, presented better training errors, with rapid convergence in all clusters. In addition, the outcomes of the evaluation of the approaches were compared in terms of the accuracy metric and presented an overall accuracy of 98.32% when using the SOM and DBN ensembles.

The encouraging outcomes of this study can be utilized practically by physicians to help them reach the right diagnosis. The research outcomes can be also deployed by scholars to explore the deployed methods on other datasets and in other medical contexts. In the context of medical disease diagnosis, there will always be room for enhancing the performance of classification models. This is particularly important in medical contexts, as the outcomes will impact human lives.

There are still plenty of future directions for researchers on dimensionality reduction, clustering, and DBN approaches for diabetes risk prediction to utilize all their advantages. Future research can deploy the same method in other contexts in the health field. The research can be also extended to assess the performance of the applied techniques on other datasets, particularly large datasets. More studies should explore this point in the design of new approaches to face the difficulties of data computation time and to benefit from big data effectively. The research on DD classification can be carried out by focusing on different stages of clustering, noise removal, and classification. Future research can deploy different methods that adopt incremental learning that adjusts the model based on the new cases. Research can follow other routes that address some issues in the datasets, such as data sparsity.

Author Contributions: Conceptualization, M.N.; methodology, M.N.; software, M.N., R.A.A., S.A., A.A. and M.A.; validation, M.N.; formal analysis, M.N.; investigation, M.N.; resources, M.N.; data curation, M.N.; writing—original draft preparation, M.N.; writing—review and editing, M.N., R.A.A., S.A., A.A. and M.A.; and visualization, M.N., R.A.A., S.A., A.A. and M.A. All authors have read and agreed to the published version of the manuscript.

Funding: The authors are thankful to the Deanship of Scientific Research at Najran University for funding this work, under the Research Groups Funding program grant code NU/RG/SERC/12/12.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data are available in the UCI machine learning archive.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. List of abbreviations utilized in this research.

Abbreviation	Description
ACC	Accuracy
ANN	Artificial Neural Network
ACO	Ant Colony Optimization
ANFIS	Adaptive Network Based Fuzzy Inference System
AIS	Artificial immune systems
CGM	Continuous Glucose Monitoring
BMI	Body Mass Index
CM	Confusion Matrix
CNN	Convolutional Neural Network
CNN-LSTM	Convolutional Neural Network Long Short-term Memory
Conv-LSTM	Convolutional Long Short-term Memory
DBN	Deep Belief Network
DD	Diabetes Disease
DDA	Diabetic Data Acquisition
DDC	Diabetic Data Classification
DIM	Diabetes Infection Measure
DM	Diabetes Mellitus
DME	Diabetic Mining and Extraction
DPDM	Diabetic Prediction and Decision Making
DT	Decision Trees
SD	Standard Deviation
FF	Farthest First
FNN	Fuzzy Neural Network
GC	Gaussian Classification
GDA	Discriminant Analysis
GP	Genetic Programming
GP-CPS	GP with comparative partner selection
IGT	Impaired Glucose Tolerance
K-NN	K-Nearest Neighbor
KAIMRCD	King Abdullah International Research Centre Diabetes
LDA-MWSVM	Morlet Wavelet Support Vector Machine Classifier
LDA	Linear Discriminant Analysis
LM	Levenberg–Marquardt
LoDI	Level of Diabetic Infection
LS-SVM	Least Squares Support Vector Machine
ML	Machine Learning

Table A1. Cont.

Abbreviation	Description
ML	Machine Learning
ML	Expectation Maximization,
MLNN	Multilayer Neural Network
NB	Naive Bayes
NN	Neural Network
NPV	Negative Predictive Value
PCA	Principle Component Analysis
PID	Pima Indian Diabetes
PK-SVM	Polynomial Kernel-Support Vector Machine
PNN	Probabilistic Neural Network
PPV	Positive Predictive Value
PR	Precision
ROC	Receiver Operating Characteristic
REC	Recall
RBFNN	Radial Basis Function Neural Networks
RNN	Recurrent Neural Network
SE	Sensitivity
SMO	Sequential Minimal Optimization
SOM	Self-Organizing Map
SP	Specificity
SVM	Support Vector Machine
SW-FFANN	Small-World Feed Forward Artificial Neural Network
TP-DL	Temporal Predictive Deep Learning
T2D	Type 2 Diabetes
T-LSTM	Traditional Long Short-term Memory
UbD	Urine-based Diabetes

References

- Mukherjee, S.; Decina, P.; Bocola, V.; Saraceni, F.; Scapicchio, P. Diabetes mellitus in schizophrenic patients. *Compr. Psychiatry* **1996**, *37*, 68–73. [[CrossRef](#)]
- Tabish, S.A. Is diabetes becoming the biggest epidemic of the twenty-first century? *Int. J. Health Sci.* **2007**, *1*, V.
- Stewart, K.J.; Dobrosielski, D.A. Exercise, Adiposity, and Regional Fat Distribution. In *Diabetes and Exercise*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 151–163.
- Pickup, J. Developing glucose sensors for in vivo use. *Trends Biotechnol.* **1993**, *11*, 285–291. [[CrossRef](#)] [[PubMed](#)]
- Ali, S.F.; Padhi, R. Optimal blood glucose regulation of diabetic patients using single network adaptive critics. *Optim. Control. Appl. Methods* **2011**, *32*, 196–214. [[CrossRef](#)]
- Newton, C.A.; Raskin, P. Diabetic ketoacidosis in type 1 and type 2 diabetes mellitus: Clinical and biochemical differences. *Arch. Intern. Med.* **2004**, *164*, 1925–1931. [[CrossRef](#)] [[PubMed](#)]
- Thunander, M.; Petersson, C.; Jonzon, K.; Fornander, J.; Ossiansson, B.; Torn, C.; Edvardsson, S.; Landin-Olsson, M. Incidence of type 1 and type 2 diabetes in adults and children in Kronoberg, Sweden. *Diabetes Res. Clin. Pract.* **2008**, *82*, 247–255. [[CrossRef](#)]
- Schäffler, A.; Büchler, C.; Müller-Ladner, U.; Herfarth, H.; Ehling, A.; Paul, G.; Schölmerich, J.; Zietz, B. Identification of variables influencing resistin serum levels in patients with type 1 and type 2 diabetes mellitus. *Horm. Metab. Res.* **2004**, *36*, 702–707. [[CrossRef](#)]
- Nayak, S.R.; Nayak, D.R.; Sinha, U.; Arora, V.; Pachori, R.B. Application of deep learning techniques for detection of COVID-19 cases using chest X-ray images: A comprehensive study. *Biomed. Signal Process. Control.* **2021**, *64*, 102365. [[CrossRef](#)]
- Alkady, W.; ElBahnasy, K.; Leiva, V.; Gad, W. Classifying COVID-19 based on amino acids encoding with machine learning algorithms. *Chemom. Intell. Lab. Syst.* **2022**, *224*, 104535. [[CrossRef](#)]
- Ahmadi, H.; Gholamzadeh, M.; Shahmoradi, L.; Nilashi, M.; Rashvand, P. Diseases diagnosis using fuzzy logic methods: A systematic and meta-analysis review. *Comput. Methods Programs Biomed.* **2018**, *161*, 145–172. [[CrossRef](#)]
- Arji, G.; Ahmadi, H.; Nilashi, M.; Rashid, T.A.; Ahmed, O.H.; Aljojo, N.; Zainol, A. Fuzzy logic approach for infectious disease diagnosis: A methodical evaluation, literature and classification. *Biocybern. Biomed. Eng.* **2019**, *39*, 937–955. [[CrossRef](#)] [[PubMed](#)]
- Ghane, M.; Ang, M.C.; Nilashi, M.; Sorooshian, S. Enhanced decision tree induction using evolutionary techniques for Parkinson's disease classification. *Biocybern. Biomed. Eng.* **2022**, *42*, 902–920. [[CrossRef](#)]
- Nilashi, M.; Abumalloh, R.A.; Minaei-Bidgoli, B.; Samad, S.; Yousoof Ismail, M.; Alhargan, A.; Abdu Zogaan, W. Predicting parkinson's disease progression: Evaluation of ensemble methods in machine learning. *J. Healthc. Eng.* **2022**, *2022*, 2793361. [[CrossRef](#)] [[PubMed](#)]

15. Nilashi, M.; Abumalloh, R.A.; Yusuf, S.Y.M.; Thi, H.H.; Alsulami, M.; Abosaq, H.; Alyami, S.; Alghamdi, A. Early diagnosis of Parkinson's disease: A combined method using deep learning and neuro-fuzzy techniques. *Comput. Biol. Chem.* **2023**, *102*, 107788. [[CrossRef](#)] [[PubMed](#)]
16. Nilashi, M.; Ahmadi, H.; Shahmoradi, L.; Ibrahim, O.; Akbari, E. A predictive method for hepatitis disease diagnosis using ensembles of neuro-fuzzy technique. *J. Infect. Public Health* **2019**, *12*, 13–20. [[CrossRef](#)]
17. Nilashi, M.; Ahmadi, H.; Sheikhtaheri, A.; Naemi, R.; Alotaibi, R.; Alarood, A.A.; Munshi, A.; Rashid, T.A.; Zhao, J. Remote tracking of Parkinson's disease progression using ensembles of deep belief network and self-organizing map. *Expert Syst. Appl.* **2020**, *159*, 113562. [[CrossRef](#)]
18. Nilashi, M.; bin Ibrahim, O.; Ahmadi, H.; Shahmoradi, L. An analytical method for diseases prediction using machine learning techniques. *Comput. Chem. Eng.* **2017**, *106*, 212–223. [[CrossRef](#)]
19. Nilashi, M.; Bin Ibrahim, O.; Mardani, A.; Ahani, A.; Jusoh, A. A soft computing approach for diabetes disease classification. *Health Inform. J.* **2018**, *24*, 379–393. [[CrossRef](#)]
20. Nilashi, M.; Ibrahim, O.; Samad, S.; Ahmadi, H.; Shahmoradi, L.; Akbari, E. An analytical method for measuring the Parkinson's disease progression: A case on a Parkinson's telemonitoring dataset. *Measurement* **2019**, *136*, 545–557. [[CrossRef](#)]
21. Kononenko, I. Machine learning for medical diagnosis: History, state of the art and perspective. *Artif. Intell. Med.* **2001**, *23*, 89–109. [[CrossRef](#)]
22. Palaniappan, R.; Sundaraj, K.; Ahamed, N.U. Machine learning in lung sound analysis: A systematic review. *Biocybern. Biomed. Eng.* **2013**, *33*, 129–135. [[CrossRef](#)]
23. Bhatia, M.; Kaur, S.; Sood, S.K.; Behal, V. Internet of things-inspired healthcare system for urine-based diabetes prediction. *Artif. Intell. Med.* **2020**, *107*, 101913. [[CrossRef](#)] [[PubMed](#)]
24. Rahman, M.; Islam, D.; Mukti, R.J.; Saha, I. A deep learning approach based on convolutional LSTM for detecting diabetes. *Comput. Biol. Chem.* **2020**, *88*, 107329. [[CrossRef](#)]
25. Zhang, K.; Chai, Y.; Yang, S.X. Self-organizing feature map for cluster analysis in multi-disease diagnosis. *Expert Syst. Appl.* **2010**, *37*, 6359–6367. [[CrossRef](#)]
26. Chen, M.C.; Chen, L.S.; Hsu, C.C.; Zeng, W.R. An information granulation based data mining approach for classifying imbalanced data. *Inf. Sci.* **2008**, *178*, 3214–3227. [[CrossRef](#)]
27. Lai, Q.; Zhou, B.; Cui, Z.; An, X.; Zhu, L.; Cao, Z.; Liu, S.; Yu, B. Development of a metabolite-based deep learning algorithm for clinical precise diagnosis of the progression of diabetic kidney disease. *Biomed. Signal Process. Control.* **2023**, *83*, 104625. [[CrossRef](#)]
28. Kannadasan, K.; Edla, D.R.; Kuppili, V. Type 2 diabetes data classification using stacked autoencoders in deep neural networks. *Clin. Epidemiol. Glob. Health* **2019**, *7*, 530–535. [[CrossRef](#)]
29. Hayashi, Y.; Yukita, S. Rule extraction using Recursive-Rule extraction algorithm with J48graft combined with sampling selection techniques for the diagnosis of type 2 diabetes mellitus in the Pima Indian dataset. *Inform. Med. Unlocked* **2016**, *2*, 92–104. [[CrossRef](#)]
30. Sisodia, D.; Sisodia, D.S. Prediction of diabetes using classification algorithms. *Procedia Comput. Sci.* **2018**, *132*, 1578–1585. [[CrossRef](#)]
31. Nilashi, M.; Ibrahim, O.; Dalvi, M.; Ahmadi, H.; Shahmoradi, L. Accuracy improvement for diabetes disease classification: A case on a public medical dataset. *Fuzzy Inf. Eng.* **2017**, *9*, 345–357. [[CrossRef](#)]
32. Amin, J.; Sharif, M.; Yasmin, M. A review on recent developments for detection of diabetic retinopathy. *Scientifica* **2016**, *2016*, 6838976. [[CrossRef](#)]
33. Polat, K.; Güneş, S.; Arslan, A. A cascade learning system for classification of diabetes disease: Generalized discriminant analysis and least square support vector machine. *Expert Syst. Appl.* **2008**, *34*, 482–487. [[CrossRef](#)]
34. Aslam, M.W.; Zhu, Z.; Nandi, A.K. Feature generation using genetic programming with comparative partner selection for diabetes classification. *Expert Syst. Appl.* **2013**, *40*, 5402–5412. [[CrossRef](#)]
35. Kahramanli, H.; Allahverdi, N. Design of a hybrid system for the diabetes and heart diseases. *Expert Syst. Appl.* **2008**, *35*, 82–89. [[CrossRef](#)]
36. ErKaymaz, O.; Ozer, M. Impact of small-world network topology on the conventional artificial neural network for the diagnosis of diabetes. *Chaos Solitons Fractals* **2016**, *83*, 178–185. [[CrossRef](#)]
37. Ganji, M.F.; Abadeh, M.S. A fuzzy classification system based on Ant Colony Optimization for diabetes disease diagnosis. *Expert Syst. Appl.* **2011**, *38*, 14650–14659. [[CrossRef](#)]
38. Dogantekin, E.; Dogantekin, A.; Avci, D.; Avci, L. An intelligent diagnosis system for diabetes on linear discriminant analysis and adaptive network based fuzzy inference system: LDA-ANFIS. *Digit. Signal Process.* **2010**, *20*, 1248–1255. [[CrossRef](#)]
39. Temurtas, H.; Yumusak, N.; Temurtas, F. A comparative study on diabetes disease diagnosis using neural networks. *Expert Syst. Appl.* **2009**, *36*, 8610–8615. [[CrossRef](#)]
40. Polat, K.; Güneş, S. An expert system approach based on principal component analysis and adaptive neuro-fuzzy inference system to diagnosis of diabetes disease. *Digit. Signal Process.* **2007**, *17*, 702–710. [[CrossRef](#)]
41. Çalışır, D.; Doğantekin, E. An automatic diabetes diagnosis system based on LDA-Wavelet Support Vector Machine Classifier. *Expert Syst. Appl.* **2011**, *38*, 8311–8315. [[CrossRef](#)]

42. Kayaer, K.; Yildirim, T. Medical diagnosis on Pima Indian diabetes using general regression neural networks. In Proceedings of the International Conference on Artificial Neural Networks and Neural Information Processing (ICANN/ICONIP), Istanbul, Turkey, 26–29 June 2003.
43. Edla, D.R.; Cheruku, R. Diabetes-finder: A bat optimized classification system for type-2 diabetes. *Procedia Comput. Sci.* **2017**, *115*, 235–242. [[CrossRef](#)]
44. Cheruku, R.; Edla, D.R.; Kuppili, V. Diabetes classification using radial basis function network by combining cluster validity index and bat optimization with novel fitness function. *Int. J. Comput. Intell. Syst.* **2017**, *10*, 247–265. [[CrossRef](#)]
45. Cruz, D.P.F.; Maia, R.D.; da Silva, L.A.; de Castro, L.N. BeeRBF: A bee-inspired data clustering approach to design RBF neural network classifiers. *Neurocomputing* **2016**, *172*, 427–437. [[CrossRef](#)]
46. Khan, N.; Gaurav, D.; Kandl, T. Performance evaluation of Levenberg-Marquardt technique in error reduction for diabetes condition classification. *Procedia Comput. Sci.* **2013**, *18*, 2629–2637. [[CrossRef](#)]
47. Maniruzzaman, M.; Kumar, N.; Abedin, M.M.; Islam, M.S.; Suri, H.S.; El-Baz, A.S.; Suri, J.S. Comparative approaches for classification of diabetes mellitus data: Machine learning paradigm. *Comput. Methods Programs Biomed.* **2017**, *152*, 23–34. [[CrossRef](#)]
48. Bozkurt, M.R.; Yurtay, N.; Yilmaz, Z.; Sertkaya, C. Comparison of different methods for determining diabetes. *Turk. J. Electr. Eng. Comput. Sci.* **2014**, *22*, 1044–1055. [[CrossRef](#)]
49. Kumari, V.A.; Chitra, R. Classification of diabetes disease using support vector machine. *Int. J. Eng. Res. Appl.* **2013**, *3*, 1797–1801.
50. Alhassan, Z.; McGough, A.S.; Alshammari, R.; Daghtani, T.; Budgen, D.; Al Moubayed, N. Type-2 diabetes mellitus diagnosis from time series clinical data using deep learning models. In *International Conference on Artificial Neural Networks*; Springer: Berlin/Heidelberg, Germany, 2018.
51. Longato, E.; Acciaroli, G.; Facchinetti, A.; Hakaste, L.; Tuomi, T.; Maran, A.; Sparacino, G. Glycaemic variability-based classification of impaired glucose tolerance vs. type 2 diabetes using continuous glucose monitoring data. *Comput. Biol. Med.* **2018**, *96*, 141–146. [[CrossRef](#)]
52. Devi, R.D.H.; Bai, A.; Nagarajan, N. A novel hybrid approach for diagnosing diabetes mellitus using farthest first and support vector machine algorithms. *Obes. Med.* **2020**, *17*, 100152. [[CrossRef](#)]
53. Kohonen, T. *Self-Organizing Maps*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2012; Volume 30.
54. Kohonen, T. Exploration of very large databases by self-organizing maps. In Proceedings of the International Conference on Neural Networks (Icnn'97), Houston, TX, USA, 9–12 June 1997; IEEE: Piscataway, NJ, USA, 1997.
55. Ahani, A.; Nilashi, M.; Yadegaridehkordi, E.; Sanzogni, L.; Tarik, A.R.; Knox, K.; Samad, S.; Ibrahim, O. Revealing customers' satisfaction and preferences through online review analysis: The case of Canary Islands hotels. *J. Retail. Consum. Serv.* **2019**, *51*, 331–343. [[CrossRef](#)]
56. Ahani, A.; Nilashi, M.; Ibrahim, O.; Sanzogni, L.; Weaven, S. Market segmentation and travel choice prediction in Spa hotels through TripAdvisor's online reviews. *Int. J. Hosp. Manag.* **2019**, *80*, 52–77. [[CrossRef](#)]
57. Ahmadi, N.; Nilashi, M.; Minaei-Bidgoli, B.; Farooque, M.; Samad, S.; Aljehane, N.O.; Zogaan, W.A.; Ahmadi, H. Eye State Identification Utilizing EEG Signals: A Combined Method Using Self-Organizing Map and Deep Belief Network. *Sci. Program.* **2022**, *2022*, 4439189. [[CrossRef](#)]
58. Nilashi, M.; Abumalloh, R.A.; Samad, S.; Minaei-Bidgoli, B.; Thi, H.H.; Alghamdi, O.A.; Ismail, M.Y.; Ahmadi, H. The impact of multi-criteria ratings in social networking sites on the performance of online recommendation agents. *Telemat. Inform.* **2023**, *76*, 101919. [[CrossRef](#)]
59. Nilashi, M.; Abumalloh, R.A.; Zibarzani, M.; Samad, S.; Zogaan, W.A.; Ismail, M.Y.; Mohd, S.; Akib, N.A.M. What Factors Influence Students Satisfaction in Massive Open Online Courses? Findings from User-Generated Content Using Educational Data Mining. *Educ. Inf. Technol.* **2022**, *27*, 9401–9435. [[CrossRef](#)]
60. Nilashi, M.; Fallahpour, A.; Wong, K.Y.; Ghabban, F. Customer satisfaction analysis and preference prediction in historic sites through electronic word of mouth. *Neural Comput. Appl.* **2022**, *34*, 13867–13881. [[CrossRef](#)]
61. Nilashi, M.; Minaei-Bidgoli, B.; Alghamdi, A.; Alrizq, M.; Alghamdi, O.; Nayer, F.K.; Aljehane, N.O.; Khosravi, A.; Mohd, S. Knowledge discovery for course choice decision in Massive Open Online Courses using machine learning approaches. *Expert Syst. Appl.* **2022**, *199*, 117092. [[CrossRef](#)]
62. Mangiameli, P.; Chen, S.K.; West, D. A comparison of SOM neural network and hierarchical clustering methods. *Eur. J. Oper. Res.* **1996**, *93*, 402–417. [[CrossRef](#)]
63. Vesanto, J.; Himberg, J.; Alhoniemi, E.; Parhankangas, J. Self-organizing map in Matlab: The SOM Toolbox. In Proceedings of the Matlab DSP Conference, Tampere, Finland, 16–17 November 1999.
64. Hinton, G.E.; Osindero, S.; Teh, Y.-W. A fast learning algorithm for deep belief nets. *Neural Comput.* **2006**, *18*, 1527–1554. [[CrossRef](#)]
65. Samui, S.; Chakrabarti, I.; Ghosh, S.K. Time–frequency masking based supervised speech enhancement framework using fuzzy deep belief network. *Appl. Soft Comput.* **2019**, *74*, 583–602. [[CrossRef](#)]
66. Erhan, D.; Courville, A.; Bengio, Y.; Vincent, P. Why does unsupervised pre-training help deep learning? In Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, Sardinia, Italy, 13–15 May 2010.
67. Chicati, M.L.; Nanni, M.R.; Cezar, E.; de Oliveira, R.B.; Chicati, M.S. Spectral classification of soils: A case study of Brazilian flooded soils. *Remote Sens. Appl. Soc. Environ.* **2017**, *6*, 39–45. [[CrossRef](#)]

68. Baltzakis, H.; Papamarkos, N. A new signature verification technique based on a two-stage neural network classifier. *Eng. Appl. Artif. Intell.* **2001**, *14*, 95–103. [[CrossRef](#)]
69. Sekhar, C.C.; Yegnanarayana, B. Neural network models for spotting stop consonant-vowel (SCV) segments in continuous speech. In Proceedings of the International Conference on Neural Networks (ICNN'96), Washington, DC, USA, 3–6 June 1996; IEEE: Piscataway, NJ, USA, 1996.
70. Jou, I.-C.; Hu, M.-S.; Juang, Y.-T. Mandarin syllables recognition based on one class one net neural network with modified selective update algorithm. In Proceedings of the 1992 IEEE International Workshop on Intelligent Signal Processing and Communication Systems (1992), San Francisco, CA, USA, 23–26 March 1992; pp. 577–591.
71. Zimmet, P.Z.; Magliano, D.J.; Herman, W.H.; Shaw, J.E. Diabetes: A 21st century challenge. *Lancet Diabetes Endocrinol.* **2014**, *2*, 56–64. [[CrossRef](#)] [[PubMed](#)]
72. Hasslachter, C.H.; Ritz, E.; Wahl, P.; Michael, C. Similar risks of nephropathy in patients with type I or type II diabetes mellitus. *Nephrol. Dial. Transplant.* **1989**, *4*, 859–863. [[CrossRef](#)]
73. Schwartz, W.B. Medicine and the computer: The promise and problems of change. In *Use and Impact of Computers in Clinical Medicine*; Springer: Berlin/Heidelberg, Germany, 1970; pp. 321–335.
74. Shamir, R.R.; Dolber, T.; Noecker, A.M.; Walter, B.L.; McIntyre, C.C. Machine learning approach to optimizing combined stimulation and medication therapies for Parkinson's disease. *Brain Stimul.* **2015**, *8*, 1025–1032. [[CrossRef](#)]
75. Sundaram, M.; Smith, M.J.; Revicki, D.A.; Elswick, B.; Miller, L.A. Rasch analysis informed the development of a classification system for a diabetes-specific preference-based measure of health. *J. Clin. Epidemiol.* **2009**, *62*, 845–856. [[CrossRef](#)] [[PubMed](#)]
76. Juniper, E.F.; Guyatt, G.H.; Jaeschke, R. How to develop and validate a new health-related quality of life instrument. In *Quality of Life and Pharmacoeconomics in Clinical Trials*, 2nd ed.; Spilker, B., Ed.; Lippincott-Raven Publishers: Philadelphia, PA, USA, 1996; pp. 49–56.
77. Hamburg, B.A.; Inoff, G.E. Relationships between behavioral factors and diabetic control in children and adolescents: A camp study. *Psychosom. Med.* **1982**, *44*, 321–339. [[CrossRef](#)]
78. Hruschka, E.R.; Ebecken, N.F. Extracting rules from multilayer perceptrons in classification problems: A clustering-based approach. *Neurocomputing* **2006**, *70*, 384–397. [[CrossRef](#)]
79. Chen, C.-H. A hybrid intelligent model of analyzing clinical breast cancer data using clustering techniques with feature selection. *Appl. Soft Comput.* **2014**, *20*, 4–14. [[CrossRef](#)]
80. Polat, K. Classification of Parkinson's disease using feature weighting method on the basis of fuzzy C-means clustering. *Int. J. Syst. Sci.* **2012**, *43*, 597–609. [[CrossRef](#)]
81. Faust, O.; Hagiwara, Y.; Hong, T.J.; Lih, O.S.; Acharya, U.R. Deep learning for healthcare applications based on physiological signals: A review. *Comput. Methods Programs Biomed.* **2018**, *161*, 1–13. [[CrossRef](#)]
82. Pham, T.; Tran, T.; Phung, D.; Venkatesh, S. Predicting healthcare trajectories from medical records: A deep learning approach. *J. Biomed. Inform.* **2017**, *69*, 218–229. [[CrossRef](#)]
83. Esteva, A.; Robicquet, A.; Ramsundar, B.; Kuleshov, V.; DePristo, M.; Chou, K.; Cui, C.; Corrado, G.; Thrun, S.; Dean, J. A guide to deep learning in healthcare. *Nat. Med.* **2019**, *25*, 24–29. [[CrossRef](#)]
84. Liu, X.Q.; Li, B.X.; Zeng, G.R.; Liu, Q.Y.; Ai, D.M. Prediction of long non-coding RNAs based on deep learning. *Genes* **2019**, *10*, 273. [[CrossRef](#)]
85. Miotto, R.; Wang, F.; Wang, S.; Jiang, X.; Dudley, J.T. Deep learning for healthcare: Review, opportunities and challenges. *Brief. Bioinform.* **2018**, *19*, 1236–1246. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.