VOL. 19, NO. 5, MARCH 2024 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

MACHINE LEARNING TECHNIQUE FOR DETECTING DIABETES DISEASE

Baydaa Hadi Saoudi[1], Tabarek Alwan Tuib[2], Yaqdhan Mahmood Hussein[2], Yusliza Yusoff [3], Fahad Taha Al-Dhief[4] and Ali Hashim Abbas[5]

[1]Technical Institute-Samawa, Information Technology and Communication (ICT), Al-Furat Al-Awast Technical University, Samawa, Iraq

[2]Collage of Engineering, Electronics and Communications Engineering, Al Muthanna University, Muthanna, Iraq

[3]Applied Industrial Analytics Research Group, Faculty of Computing, Universiti Teknologi Malaysia, Johor Bahru, Johor, Malaysia

[4]Faculty of Engineering, School of Electrical Engineering, Universiti Teknologi Malaysia (UTM), Johor Bahru, Malaysia

[5]Department of Computer Technical Engineering, College of Information Technology, Imam Ja'afar Al-Sadiq University, Al-Muthanna, Iraq

E-Mail: fahadtaha246@gmail.com

ABSTRACT

Diabetes is considered a critical disease and it has been a growing concern owing to its increased morbidity. Moreover, the average age of people who are affected by diabetes illness has currently declined to the mid-20s. Given the high prevalence, it is necessary to address this problem effectively. Due to the significant prevalence of diabetes illness, it is essential to handle and address this issue appropriately. Currently, machine learning methods are considered a vital area for detecting and diagnosing disease. These methods can learn from data and classify data based on the coordinate subjects. This paper presents a model for detecting diabetes illness based on a machine learning technique. The Support Vector Machine (SVM) algorithm is used for classifying the people who are categorized as patients with diabetes disease from the people who are categorized as non-diabetic. Further, the database is compiled from the Pima Indian Diabetes Dataset (PIDD). The results show that the proposed model achieves 81.8% accuracy. Moreover, the proposed model achieves 84.34% sensitivity and 74.35% specificity.

Keywords: machine learning, SVM, healthcare, diabetes disease, PIDD database.

Manuscript Received 12 December 2023; Revised 12 March 2024; Published 30 April 2024

1. INTRODUCTION

One of the most important areas in the field of research with the purpose of detecting or predicting subjects is machine learning [1]. The algorithms of machine learning are applied in various domains of healthcare such as voice pathology detection [2], federated learning based on fog-cloud [3], COVID-19 detection [4], cancer detection [5], autism detection [6], and diabetic detection [7]. According to World Health Organization statistics, India has the highest prevalence of diabetes [8]. Furthermore, The overall number of diabetics in India in the year 2000 was 31.7 million, and it is expected to rise to 79.4 million by 2030 [9]. Nevertheless, diabetes is considered a chronic disease condition that is characterized by increased blood sugar levels over an extended period. In some cases, there is organ failure, including the heart, liver, stomach, kidneys, and others, where these organs are failed induced in the long term because of diabetes's influence [10]. Consequently, the detection of this disease by using an efficient technique and non-invasive process is crucial.

On the other hand, there are several Artificial Intelligence (AI) approaches that have been employed for diabetes detection because they can mine data and learn from it to deliver improved results [11]. The AI allows computers to think [12]. Further, the AI enhances the intelligence of computers. Machine learning is a branch of AI research. Some researchers believe that intelligence cannot be developed without learning. Figure 1 depicts many different kinds of machine learning methods [13]. All these methods are employed to classify the subjects of the dataset. In supervised learning, a training set of examples with appropriate objectives is provided, and algorithms react appropriately to all possible inputs based on this training set. Supervised Learning includes classification and regression [14]. In addition, appropriate responses or goals are not offered in unsupervised learning. The unsupervised learning approach attempts to discover similarities between input data and based on such similarities, then it can classify the data [15]. A semi-supervised learning approach is considered one of the supervised learning types. Furthermore, the unlabeled data is utilized for the training task in this approach [16]. In reinforcement learning, behaviorist psychology promotes this kind of learning. The algorithm is notified when the answer is incorrect, but it is not advised on how to rectify it. It must investigate and test numerous alternatives until it gets the correct answer [17]. In evolutionary learning, it applies evolutionary methods to address optimization issues in machine learning, and these methods have shown promising results in a variety of applications [18]. Moreover, deep learning is a subset of machine learning that is built on a collection of algorithms [19]. These learning algorithms in data represent high-level abstraction.

As mentioned previously, diabetes is a serious disease, and detecting it in an early stage is critical to avoiding catastrophic side effects. The most prevalent

symptoms of diabetes disease are improper metabolism, hyperglycemia, and an increased risk of particular problems affecting the kidneys, eyes, and nervous system. All these parts are considered vital organs in the human body. Additionally, these symptoms are utilized to collect data, and then modelling is done based on gender and age categories [20]. However, current medical practice requires that a patient undergo a battery of tests to gather the information needed for diagnosis, following which therapy is delivered depending on the diagnosis [21]. Machine learning approaches are desperately needed in today's environment to reduce human effort while also achieving greater automation with fewer errors [22]. Therefore, this paper presents a model based on a machine learning technique for detecting diabetes illness. In addition, the technique of machine learning is learned based on a diabetes database and evaluated to find out its performance in the identification of diabetes illness. The proposed machine learning technique is named Support Vector Machine (SVM) and it is evaluated and compared with other algorithms in the detection of diabetes disease.
Moreover, this paper is organized as follows: Section 2 presents the related work. Section 3 gives the proposed method. Section 4 presents the diabetic database. Section 5 presents the classification algorithm. Section 6 discusses the experimental results of the proposed method. Lastly, Section 7 gives the conclusion of this paper.

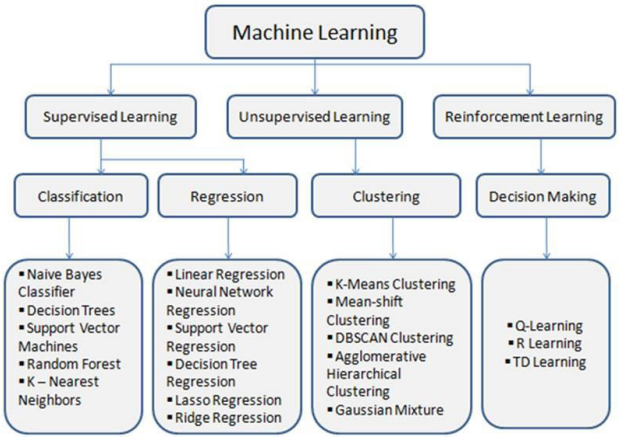

Figure-1. Types of machine learning methods.

2. RELATED WORK

The investigation of related work provides outcomes on numerous healthcare datasets, where assessments and predictions were performed using different approaches and techniques. Many researchers have designed and implemented numerous prediction models employing different types of machine learning algorithms, data mining methods, or the incorporation of these approaches [23]. The study in [24] has proposed different classification techniques for diabetes prediction based on machine learning utilizing various classifiers. In other words, four classifier models are employed in this study which are Bagging, Random Forest (RF), Gradient Boosting (GB), and AdaBoosting. The diabetes prediction system using those classifiers constructed for this study is measured by accuracy, precision, F1-score, and recall. The accuracy of the RF model is up 75%, while the accuracy of the Bagging, AdaBoosting, and Gradient Boosting classifiers is 72%, 71%, and 76%, respectively. With a prediction accuracy of 76% for the Gradient Boosting classifier, it can be concluded that the Gradient Boosting classifier has achieved the best prediction accuracy among all ensemble classifiers tested to predict diabetes.

In addition, a new framework is presented to study and evaluate diabetes prediction models based on machine learning techniques, namely K-Nearest Neighbor (KNN) and RF [25]. These techniques are regarded as the most widely used in the prediction part. Furthermore, in this study, it is suggested that a novel intelligent diabetes mellitus prediction framework be constructed utilizing machine learning. The authors presented the training processes, model evaluation strategies, and challenges related to diabetes prediction, as well as they provided solutions using the presented framework. The presented study achieved 83% prediction accuracy that has been obtained by the RF technique. Moreover, the KNN technique achieved 65% prediction accuracy. The results of this study may be useful to stakeholders, researchers, health professionals, and students who are working in diabetes prediction research.

This study in [26] introduced two methods for diabetes prediction. The first is a classification-based method, whereas the second is a hybrid algorithm. Also, The RF technique is used for the identification. The XGBoost technique was used for the hybrid approach. These two methods were developed and compared to investigate the prediction accuracy in diabetes for two different machine learning techniques. The authors have utilized the PIMA diabetes database that has been presented by the National Institute of Diabetes (NID). The presented method obtained an accuracy score of 74.10% that has been obtained by the XGBoost algorithm, which is higher than the RF technique. Moreover, the RF algorithm achieved 71.9% accuracy.

In the work in [27], different deep learning methods are presented to predict diabetes persons. Furthermore, it involves feature augmentation with a sparse autoencoder (SAE), data augmentation with a variational autoencoder (VAE), and a Convolutional Neural Network (CNN) which has been used for the classification. The Pima Indians Diabetes Database (PIDD) is used for training and testing the proposed methods, where this database considers patient information that needs to be analyzed. For example, insulin level (i.e., glucose), age, blood pressure, and number of pregnancies. The outcomes of this study show that the presented method has achieved 92.31% detection accuracy which has been obtained by the CNN technique.

Additionally, the assessment of the PIDD dataset has been presented also in [28]. The standard deviation of characteristics is employed in this work. In the first stage, it has been used the standard deviation of attributes as power for computing K-Nearest Neighbour (KNN) to enhance classification accuracy. While, in the second stage, distance in KNN is processed based on the mean of

the standard deviation of attributes to further enhance the classification accuracy. The analysis of the PIDD database is conducted by dividing the database into 90% training data and 10% testing data. The outcomes of this method showed that the average classification accuracy was 83.2%, which is a significant increase over other standard techniques.

Deep learning technique is used and presented for diabetic eye disease identification in [29]. The goal of this study was to develop an automated classification system for two scenarios: a) multi-class Diabetic Eye Illness (DEI) and b) Moderate Multi-Class (MMC). The presented model was evaluated on a variety of datasets that had been annotated by an ophthalmologist. The experiment was carried out using CNN models. Additionally, several performance enhancement approaches are used, such as fine-tuning, contrast enhancement, and optimization. The VGG16 model achieved a maximum accuracy of 85.95% for moderate multi-class classification and 88.3% for multi-class classification. According to the previously mentioned studies, Table-1 summarizes works that are related to diabetes detection using the algorithms of machine learning and deep learning.

Table-1. The summary of related works for diabetes detection.

Years	Techniques	Databases	Results	Reference
2023	RF, GB, Bagging, and AdaBoosting	PIMA Indian Diabetes Dataset	RF = 75% GB = 76% Bagging = 72% AdaBoosting = 71%	[24]
2022	KNN and RF	Pima Indian Diabetes Database (PIDD)	RF = 83% KNN = 65%	[25]
2021	RF and XGBoost	PIMA	RF = 71.9% XGBoost =74.10%	[26]
2021	CNN	PIDD	92.31%	[27]
2021	KNN	PIDD	83.2%	[28]
2020	CNN	Messidor	85.95% and 88.3%	[29]

3. PROPOSED METHOD

The proposed work aims to detect diabetes illness by using a machine learning algorithm. The proposed method consists of two main stages, where the first stage indicates the diabetes database. Furthermore, the second stage refers to the algorithm to classify the abnormal class (i.e., diabetes) from the normal class. In addition, Figure-2 shows the flowchart of the proposed method. These two stages will be explained in the following sections, respectively.

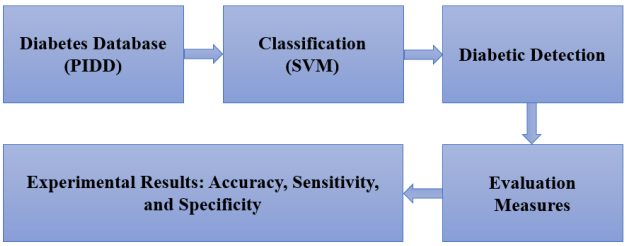

Figure-2. The flowchart of the proposed method.

4. DIABETES DATASET

The proposed model is assessed and tested using the Pima Indian Diabetes Dataset (PIDD) [30], which is considered an open-source database and it has been gathered from the Kaggle Repository [31]. The dataset contains numerous medical predictor variables such as the number of pregnancies the patient has had, insulin level, BMI (Body Mass Index), and age. Besides, it has one target variable which refers to the outcome. The age of users in the dataset is in the range of 21-81. In other words, the age for the youngest volunteer is 21 years old and the age for the oldest user is 81 years old. More details of this dataset can be found in [31].

Furthermore, this dataset contains 8 attributes (Features) and two class outputs (0 and 1), where class 0 represents a patient without diabetes and class 1 represents a patient with diabetes. In the PIDD database, there are 768 cases, 268 cases are referred to class 1, which is considered diabetic patients, and 500 cases are referred to class 0, which is considered non-diabetic patients. In addition, the PIDD database in this work is divided into 80% and 20% for training and testing the proposed model, respectively.

5. CLASSIFICATION ALGORITHM

The Support Vector Machine (SVM) is considered a high-performing machine learning method that is used for classification intention [32]. In addition, it is one of the learned model techniques for predicting or classifying unknown data. The data is plotted in n-dimensional space in the SVM approach, with the data acting as the point in the plot. Moreover, the data is plotted according to its category collected as a set in the graph. In other words, the SVM algorithm can be utilized to find the hyperplane that separates the data into two categories or classes. Figure-3 shows the SVM algorithm for two-dimensional data.

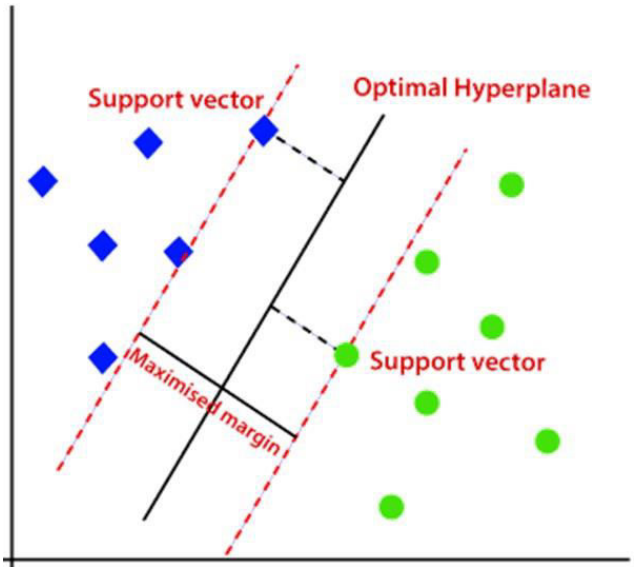

Figure-3. The SVM algorithm.

The SVM approach involves separating or dividing the dataset into two groups. The first group indicates the training process. In this group, the training process is conducting the training step by constructing a model based on the training data group or known data that has been acquired from the experimentation. Whilst, the second group indicated the testing process. In this group, the testing process conducts the classification of unknown data based on the training model. Thereafter, the testing group discovers the suitable class based on the coordinate range. In the SVM algorithm, every instance in the training data contains target values as well as different features or attributes. The fundamental goal of the SVM algorithm is to create a model that predicts the target values of the test data based on its features using the training data [33]. However, the same as other machine learning algorithms, the SVM algorithm performs poorly on noisy datasets with overlapping classes [34]. The SVM algorithm is utilized in systems for text identification [35], intrusion detection [36], and image processing [37]. Also, the SVM algorithm is easier than neural networks in terms of implementation [38]. Moreover, the SVM tool is considered an effective algorithm, particularly in high-dimensional space [39].

6. RESULTS DISCUSSIONS

The proposed model of machine learning is presented for the detection of diabetes disease. The proposed model uses the SVM algorithm for classifying the instances of diabetes patients' class from the instances of non-diabetes people class. Moreover, the instances of both classes have been gathered from the PIDD database, which was created in India. From this database, there are 268 cases are referred to class 1, which considered diabetic patients, and 500 cases are referred to class 0, which considered non-diabetic patients.

Accordingly, the total number of instances used in the proposed model is 768 cases. Besides, in this study, the PIDD database is split into 80% and 20% for training and testing the proposed model, respectively. The proposed model is carried out based on Python over PC (windows 10 Home Edition 64-bit OS), Intel Core (TM) CPU i5-5200U, @2.30GHz, and 16GB RAM. For evaluating the performance of the proposed model, we have used Accuracy (ACC), Sensitivity (SEN), and Specificity (SPE) as performance measures. These measures are computed as the following equations [40]:

$$ACC = (TP + TN)/(TP + TN + FP + FN) \quad (1)$$

$$SEN = TP/(TP + FN) \quad (2)$$

$$SPE = TN/(TN + FP) \quad (3)$$

Based on these measures, the performance of the proposed model is evaluated. Figure-4 shows the performance results of the proposed model in the detection of diabetes disease in terms of ACC, SEN, and SPE measures. The outcomes have shown that the proposed model based on the SVM algorithm achieved 81.8% ACC. Further, the performance of the SVM algorithm gained 84.34% SEN and 74.35% SPE.

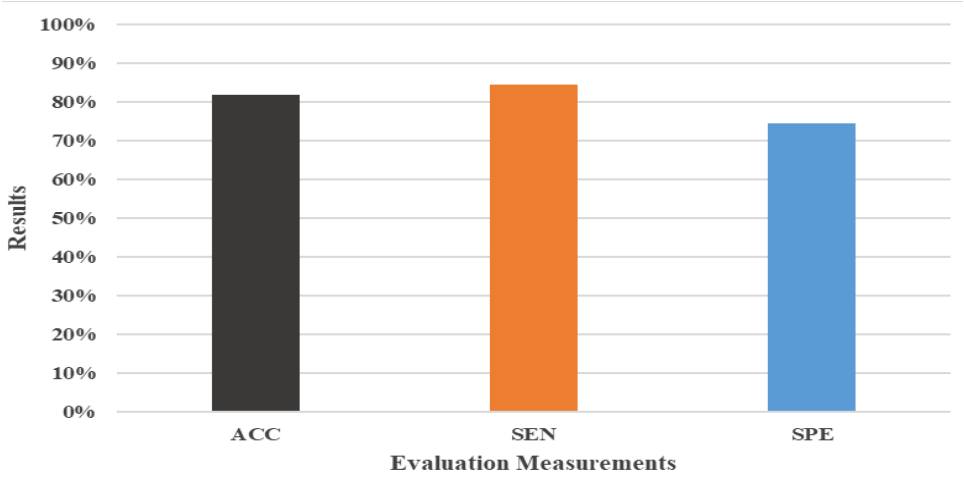

Figure-4. The performance results of the proposed model.

Furthermore, Table-2 exhibits the confusion matrix of detection results for the diabetes disease utilizing the SVM algorithm. According to the obtained results, the performance of the proposed model based on the SVM algorithm showed satisfactory outcomes in the detection of diabetes disease.

Table-2. The confusion matrix.

	Predicted Class	
Actual Class	97	10
	18	29

In addition, the performance of the proposed model is evaluated with other works for the detection of diabetes disease such as the work in [41] is used the RF algorithm, Deep Neural Network (DNN) [42], Naïve Bayes (NB) [43], and Decision Tree (DT) [44]. The results showed that the performance of the proposed model based on the SVM algorithm outperformed the other methods in the detection of diabetes disease in terms of detection accuracy. Table 3 shows the comparison in terms of the accuracy between the proposed model and other methods.

Table-3. Comparison of accuracy between methods.

Methods	**Accuracies**
Proposed SVM	81.8%
RF [41]	79.57
DNN [42]	77.09%
NB [43]	79%
DT [44]	72%

7. CONCLUSIONS

Machin learning methods have been used effectively in various areas, including medical diagnostics. In this study, a model for detecting diabetes illness is presented. The proposed model is based on the SVM algorithm that has been used for classifying the people who are categorized as non-diabetic (class 0) from the people who are categorized as patients with diabetes (class 1). The instances of both classes are gathered from the PIDD database. In class 0, there were 500 instances, while in class 1, there were 268 instances. Furthermore, the performance of the proposed model is assessed by ACC, SEN, and SPE. The outcomes of the proposed model which is based on the SVM algorithm have shown that the obtained ACC was 81.8%. Whereas, the results of the SEN and SPE were 84.34% and 74.35%, respectively. The performance of the proposed model can be considered adequate in diabetes detection. Moreover, the performance of the proposed model can be improved in future work by tuning the parameters of the SVM algorithm. Another future work can be assessing the proposed model with a varied number of instances or a balanced database.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the financial support from the Ministry of Higher Education Malaysia (MOHE) under the Fundamental Research Grant Scheme (FRGS) (Ref: FRGS/1/2021/ICT07/UTM/02/5) vote R.J130000.7851.5F462 and Research Management Center (RMC) of Universiti Teknologi Malaysia (UTM).

REFERENCES

[1] M. A. A. Albadr, S. Tiun, M. Ayob, F. T. AL-Dhief, K. Omar and M. K. Maen. 2022. Speech emotion recognition using optimized genetic algorithm-extreme learning machine. Multimedia Tools and Applications. 81(17): 23963-23989.

[2] F. T. AL-Dhief et al. 2020. Voice pathology detection using machine learning technique. in 2020 IEEE 5th International Symposium on Telecommunication Technologies (ISTT), IEEE. pp. 99-104.

[3] A. Lakhan, M. A. Mohammed, S. Kadry, K. H. Abdulkareem, F. T. Al-Dhief and C.-H. Hsu. 2021. Federated learning enables intelligent reflecting surface in fog-cloud enabled cellular network. PeerJ Computer Science. 7: 758.

[4] M. A. A. Albadr, S. Tiun, M. Ayob and F. T. Al-Dhief. 2022. Particle swarm optimization-based extreme learning machine for COVID-19 detection. Cognitive Computation. pp.1-16.

[5] S. Ubaidillah, R. Sallehuddin and N. A. Ali. 2013. Cancer detection using artificial neural network and support vector machine: a comparative study. Jurnal Teknologi. 65(1): 73-81.

[6] R. N. S. Husna, A. Syafeeza, N. A. Hamid, Y. Wong and R. A. Raihan. 2021. Functional magnetic resonance imaging for autism spectrum disorder detection using deep learning Jurnal Teknologi. 83(3): 45-52.

[7] M. A. A. Albadr, M. Ayob, S. Tiun, F. T. Al-Dhief and M. K. Hasan. 2022. Gray wolf optimization-extreme learning machine approach for diabetic retinopathy detection. Frontiers in Public Health. 10: 925901.

[8] V. Mohan, S. Sandeep, R. Deepa, B. Shah and C. Varghese. 2007. Epidemiology of type 2 diabetes: Indian scenario. Indian journal of medical research. 125(3): 217-230.

[9] S. A. Kaveeshwar and J. Cornwall. 2014. The current state of diabetes mellitus in India. The Australasian Medical Journal. 7(1): 45.

[10] K. Sumangali, B. Geetika and H. Ambarkar. 2016. A classifier based approach for early detection of diabetes mellitus. in 2016 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), IEEE, pp. 389-392.

[11] S. Gujral. 2017. Early diabetes detection using machine learning: a review. ed: IJIRST.

[12] A. A. Barreda-Luna, J. Rodríguez-Reséndiz, O. Rodríguez-Abreo and J. M. Álvarez-Alvarado. 2022. Spatial Models and Neural Network for Identifying Sustainable Transportation Projects with Study Case in Querétaro, an Intermediate Mexican City. Sustainability. 14(13): 7796.

[13] M. Fatima and M. Pasha. 2017. Survey of machine learning algorithms for disease diagnostic. Journal of Intelligent Learning Systems and Applications. 9(01): 1-16.

[14] Z.-H. Zhou. 2018. A brief introduction to weakly supervised learning. National Science Review. 5(1): 44-53.

[15] F. T. Al-Dhief *et al*. 2020. A survey of voice pathology surveillance systems based on the Internet of Things and machine learning algorithms. IEEE Access. 8: 64514-64533.

[16] X. Yang, Z. Song, I. King and Z. Xu. 2022. A survey on deep semi-supervised learning. IEEE Transactions on Knowledge and Data Engineering.

[17] M. Laskin, K. Lee, A. Stooke, L. Pinto, P. Abbeel and A. Srinivas. 2020. Reinforcement learning with augmented data. Advances in neural information processing systems. 33: 19884-19895.

[18] H. Al-Sahaf *et al*. 2019. A survey on evolutionary machine learning. Journal of the Royal Society of New Zealand. 49(2): 205-228.

[19] S. Dong, P. Wang and K. Abbas. 2021. A survey on deep learning and its applications. Computer Science Review. 40: 100379.

[20] T. Sharma and M. Shah. 2021. A comprehensive review of machine learning techniques on diabetes detection. Visual Computing for Industry, Biomedicine, and Art. 4: 1-16.

[21] N. D'Souza, K. Shah, and P. Singh. 2022. Diabetes Detection Using Machine Learning Algorithms. in 2022 IEEE Bombay Section Signature Conference (IBSSC), IEEE. pp. 1-5.

[22] K. Jha, A. Doshi, P. Patel, and M. Shah. 2019. A comprehensive review on automation in agriculture using artificial intelligence. Artificial Intelligence in Agriculture. 2: 1-12.

[23] A. Mujumdar and V. Vaidehi. 2019. Diabetes prediction using machine learning algorithms. Procedia Computer Science. 165: 292-299.

[24] M. Bhattacharya and D. Datta. 2023. Development of Predictive Models of Diabetes Using Ensemble Machine Learning Classifier. in International Conference on Advancements in Smart Computing and Information Security: Springer, pp. 377-388.

[25] R. Krishnamoorthi *et al*. 2022. A novel diabetes healthcare disease prediction framework using machine learning techniques. Journal of Healthcare Engineering. Vol. 2022.

[26] S. Barik, S. Mohanty, S. Mohanty and D. Singh. 2021. Analysis of prediction accuracy of diabetes using classifier and hybrid machine learning techniques. in Intelligent and Cloud Computing: Proceedings of ICICC 2019, 2: Springer, pp. 399-409.

[27] M. T. García-Ordás, C. Benavides, J. A. Benítez-Andrades, H. Alaiz-Moretón and I. García-Rodríguez. 2021. Diabetes detection using deep learning techniques with oversampling and feature augmentation. Computer Methods and Programs in Biomedicine. 202: 105968.

[28] R. Patra. 2021. Analysis and prediction of Pima Indian Diabetes Dataset using SDKNN classifier technique. in IOP Conference Series: Materials Science and Engineering, 1070(1): IOP Publishing, p. 012059.

[29] R. Sarki, K. Ahmed, H. Wang, and Y. Zhang. 2020. Automated detection of mild and multi-class diabetic eye diseases using deep learning. Health Information Science and Systems. 8(1): 32.

[30] J. W. Smith, J. E. Everhart, W. Dickson, W. C. Knowler and R. S. Johannes. 1988. Using the ADAP

learning algorithm to forecast the onset of diabetes mellitus. inroceedings of the annual symposium on computer application in medical care: American Medical Informatics Association. p. 261.

[31] U. M. Learning. Pima Indians Diabetes Database. Kaggle Repository. https://www.kaggle.com/datasets/uciml/pima-indians-diabetes-database?resource=download (accessed.

[32] F. T. AL-Dhief, N. M. a. A. Latiff, M. M. Baki, N. N. N. A. Malik, N. Sabri and M. A. A. Albadr. 2021. Voice pathology detection using support vector machine based on different number of voice signals. in 2021 26th IEEE Asia-Pacific Conference on Communications (APCC), 2021: IEEE, pp. 1-6.

[33] X. Zhou et al. 2020. Data-driven thermal comfort model via support vector machine algorithms: Insights from ASHRAE RP-884 database. Energy and Buildings. 211: 109795.

[34] B. S. Bhati and C. S. Rai. 2020. Analysis of support vector machine-based intrusion detection techniques. Arabian Journal for Science and Engineering. 45: 2371-2383.

[35] S. Chatterjee, P. G. Jose and D. Datta. 2019. Text classification using SVM enhanced by multithreading and CUDA. International Journal of Modern Education and Computer Science. 12(1): 11.

[36] M. Mohammadi *et al.* 2021. A comprehensive survey and taxonomy of the SVM-based intrusion detection systems. Journal of Network and Computer Applications. 178: 102983.

[37] H. I. Dino and M. B. Abdulrazzaq. 2019. Facial expression classification based on SVM, KNN and MLP classifiers. in 2019 International Conference on Advanced Science and Engineering (ICOASE), 2019: IEEE. pp. 70-75.

[38] G. Alfonso-Francia *et al*. 2022. Performance Evaluation of Different Object Detection Models for the Segmentation of Optical Cups and Discs. Diagnostics. 12(12): 3031.

[39] S. M. Erfani, S. Rajasegarar, S. Karunasekera and C. Leckie. 2016. High-dimensional and large-scale anomaly detection using a linear one-class SVM with deep learning. Pattern Recognition. 58: 121-134.

[40] M. A. A. Albadr, S. Tiun, M. Ayob, F. T. Al-Dhief, T.-A. N. Abdali and A. F. Abbas. 2021. Extreme learning machine for automatic language identification utilizing emotion speech data. In 2021 International Conference on Electrical, Communication, and Computer Engineering (ICECCE), 2021: IEEE, pp. 1-6.

[41] V. Chang, J. Bailey, Q. A. Xu and Z. Sun. 2023. Pima Indians diabetes mellitus classification based on machine learning (ML) algorithms. Neural Computing and Applications. 35(22): 16157-16173.

[42] S. A. Alex, J. J. V. Nayahi, H. Shine and V. Gopirekha. 2022. Deep convolutional neural network for diabetes mellitus prediction. Neural Computing and Applications. 34(2): 1319-1327.

[43] F. A. Khaleel and A. M. Al-Bakry. 2022. Diagnosis of diabetes using machine learning algorithms. Materials Today: Proceedings. 80: 3200-3203.

[44] S. Shafi and G. A. Ansari. 2021. Early prediction of diabetes disease & classification of algorithms using machine learning approach. in Proceedings of the International Conference on Smart Data Intelligence (ICSMDI 2021).