2026 Volume: 02 Issue: 09 www.ijrpa.com ISSN 2456-9995 Research Article

International Journal Research Publication Analysis

Page: 01-14

DETECTION OF THE LEVELS OF DIABETES MELLITUS IN HUMAN BEINGS USING DEEP LEARNING MODEL

***Dr. C. Premila Rosy**

Head & Assistant Professor, PG & Research Department of Computer Science, Idhaya College for Women, Affiliated to Bharathidasan University, Kumbakonam.

Article Received: 06 August 2026
Article Revised: 26 August 2026
Published on: 16 September 2026

***Corresponding Author: Dr. C. Premila Rosy**
Head & Assistant Professor, PG & Research Department of Computer Science, Idhaya College for Women, Affiliated to Bharathidasan University, Kumbakonam.
DOI: https://doi-doi.org/101555/ijrpa.7416

ABSTRACT

The diabetes mellitus is a prominent unending rapture to the human civilization that affects the health of individuals and causes unending health hazards. The automatic diagnostic models offers a greater source and detection and prediction of diabetes mellitus that helps clinician to identify the level of diabetes in humans. However, most of the methods fail in detecting the levels diabetes mellitus due to its complexity in estimating the levels. In this paper, the study considers various parameter of human body to find the level of diabetes mellitus. To achieve this, a deep learning paradigm is used that helps in feature extraction, selection and classification of instances in predicting the levels of diabetes mellitus. The simulation is conducted to test the efficacy of the model, and the results show that the deep learning paradigm provides increased accuracy, specificity, sensitivity and reduced mean absolute percentage error than other models.

KEYWORDS: Diabetes Mellitus, Data Mining, Prediction, Level of Diabetes.

1. INTRODUCTION

Significant advancement in biotechnology, and in particular high-performance sequencing, resulted in easy and efficient information collection, which included applied biological research in the big data sector [1] [2].

To date, in addition to high-precision sequencing methods, a range of digital machinery and sensors in a multitude of areas, including super-resolution digital microscopy, magnetic resonance imaging (MRI). Although these technologies produce different results, they have not been analysed, interpreted or extracted. In this regard, working in the field of biological data mining or data analysis of biological data is more than ever appropriate and significant. The main goal is to explore the increasingly increasing collection of biological data and to lay the groundwork for responses to particular biological and medical issues.

The strategies are effective and successful due to the ability, through measured methods, to draw patterns and to model data. This is particularly important in the large data age, especially where terabytes or petabytes of data can be greater than the dataset. Data access has also significantly enhanced biological research based on data. One of the most important research applications is pronostic and diagnostic in this hybrid market along with human threats and/or disease avoidance in terms of quality of life.

One such disease is diabetes mellitus (DM). Data mining methods are a critical method for the use of large quantities of knowledge about diabetes in the exploitation of DM scientific expertise. DM is one of the core medical research priorities which inevitably provides vast amounts of data, taking account of the extreme social impact of each disease. Diabetes experiments, including machine learning and extraction of data, have been classified in a substantial way.

Our aim was to produce a DSVM solution that identified people without diabetes or pre-diabetes without being diagnosed. The DSVM classification is compared in deep learning models with predictions from the same range of features.

2. Background

In resolving problems in the classification of many biomedical fields, in particular in bioinformatics, the SVM algorithm has demonstrated high performance [3] with supervised machine learning. Rather, the SVM separates two groups by building a hyperplane that distinguishes classes optimally after the input data is converted mathematically from a logistic regression, which depends on a predetermined model to forecast events in a binary case or not. Since the SVM approach is data-driven and model-free, there may be considerable discrimination in classification capabilities of the SVM approach, especially where space is large. This technology was recently used to automatically identify diseases and to enhance methods for the diagnosis of clinical diseases [4]. [5]. [5].

In assessing the SVM's ability to classify individuals into disease-specified groups, we have chosen diabetes as an example. In the USA there are an estimated 23.6 million people who have diabetes, about a 3rd of whom do not know that they are affected [6]. 57 million additional people suffer from prediabetes, which increases their risk of diabetes, heart disease and stroke with higher blood glucose levels. Recent findings indicate that diabetes may be prevented in people with prediabetes through diets or pharmacotherapy [7-9].

Early diagnosis and screening is also key to effective prevention measures [10]. Several risk rates and estimation equal treatments have been developed [11]-[13] to distinguish people with higher diabetes or pre-diabetes risks based on general risk factors such as the Body Mass Index (BMI) and family history of diabetes.

For instance, logistic regressions in a newly released risk computer are used to separate pre-diabetes and diabetes patients using combinations of different risk variables [14]. Our aim was to develop an SMV method to distinguish between people without either diabetes or pre-diabetes from people without those conditions. The variables to make the SVM models were limited to simple therapeutic procedures which were laboratory free. The logistic regression models with the same set of variables are contrasted with these system forecasts. The ultimate goal was to create a web classification system to demonstrate the SVM approach [15]-[22].

3. Proposed Method

Knowledge Discovery Database in the proposed model involves the attempt to pre-process the input data, extraction and selection of features in training and testing phase, and finally the classification of instances using Deep SVM. The architecture of which is given in Figure 1.

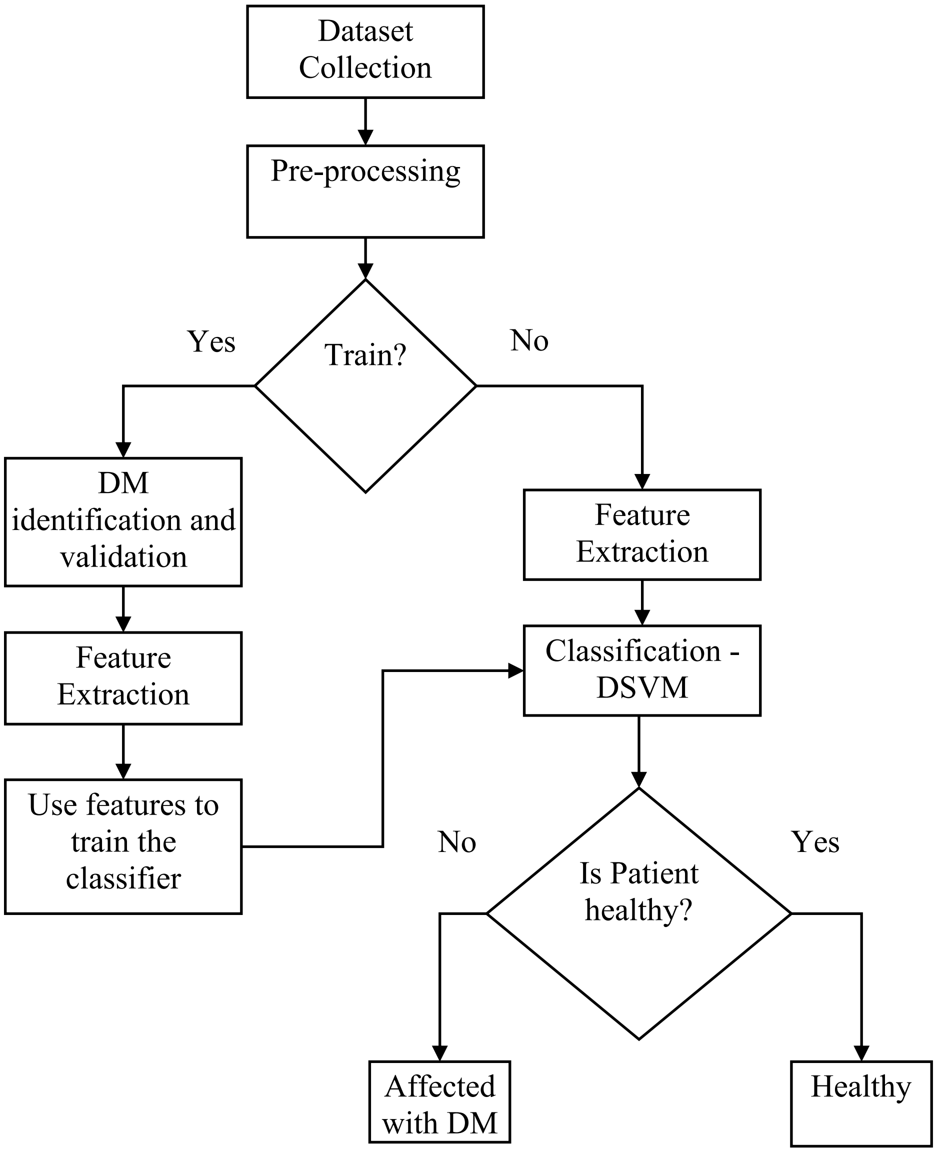

Figure 2: Proposed Method.

Classification

SVM is a supervised methodology used to identify and discern patterns systematically. The SVM algorithm classifies the distance between two clusters by creating a hyperplane that divides optimally between two sets. The use of special nonlinear functionality kernels to turn this algorithm into multidimensional space in the input space achieves a powerful discriminatory ability.

The basic idea behind the DSVM technique is to create n-1 dimension hyperplanes to distinguish two classes in an n-dimensional space.A data collection is an N-dimensional Vector. For more dimensions, DSVM strives to achieve the ideal hyperplane separation specified as the biggest margin. It maximizes the distance between the hyperplane and the nearest point, known as support vectors. The easiest way to separate two groups by a linear hyperplane. However, real situations aren't always that simple. In the two classes, certain data points that do not readily differentiate in a gray field. DSVM resolves this by inserting a user-specified parameter "C" which specifies the balance between minimizing misclassification and maximizing the margin; using kernel functions, with some data points on the wrong side. The DSVM process tries to give organizations a meaningful difference from multiple logistic regressions without including an estimate of the probabilities of class inclusion in the data set.

In this paper we suggest that the multi-class SVM objective be followed up by deep neural nets for classification tasks. The weights of the lower layer are known by replicating the gradients from the SVM. We need to distinguish the SVM objective from the triggering of the pre-last sheet.

Consider $l(w)$ as the objective function, where the input x is considered as the activation function h,

$$\frac{\partial l(w)}{\partial h_n} = -Ct_n\left(I\left(l > w^T h_n t_n\right)\right) \qquad (1)$$

Where

I{·} – activation function.

For the DSVM, we have

$$\frac{\partial l(w)}{\partial h_n} = -2Ct_n w\left(\max\left(l - w^T h_n t_n, 0\right)\right) \qquad (2)$$

Here, the backpropagation neural network (BPNN) is employed instead of softmax deep learning networks.

4. RESULTS AND DISCUSSIONS

This section covers the assessment of the high-end computation classification template proposed by the Matlab method. Test data sets were used for analysis of model performance. Due to the overfitting of test data sets, the validation of test data sets has avoided potential performance prediction biases. A logistic regression feature built during the training phase using the logistic regression model calculated the prediction value for each member of the test data collection.

On the training data set for the robustness of the predictions on the DSVM models, 10-fold cross-validations were performed. The data set was divided into 10 similarly large bits. The test data set for each sub-set was used to train a model in each case and randomly selected in the same number of cases from nine remaining sub-dates. This procedure has been repeated ten times and can be used for test outcomes once in each sub-set.

The DSVM classifier is tested against accuracy, sensitivity, specificity and f-measure. The DBN framework is compared against back propagation neural network (BPNN), feed forward neural network (FFNN), Long Short Term Memory (LSTM) and Recurrent Neural Network (RNN) to validate the accuracy of the classifier.

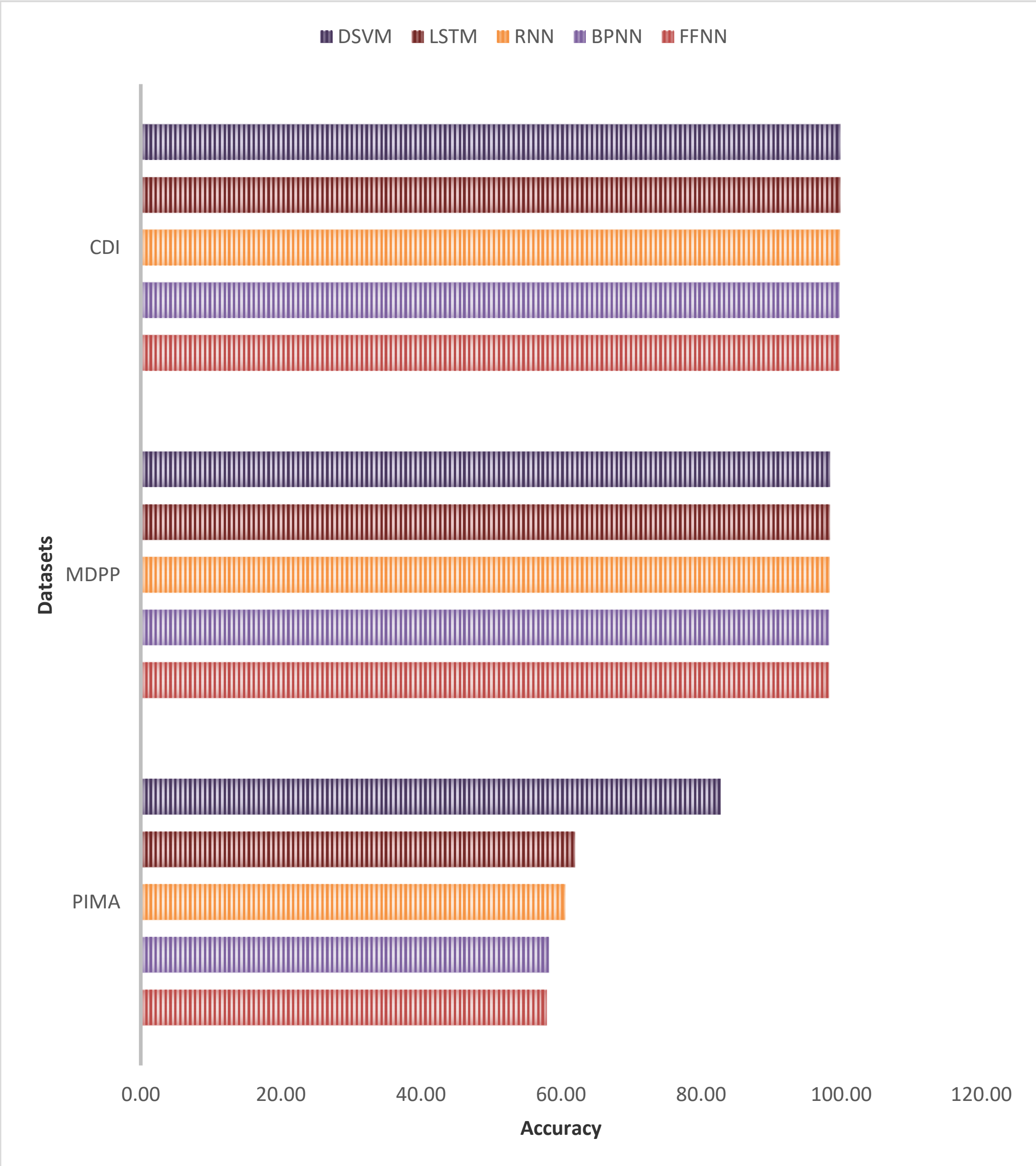

Figure 3. Accuracy with various datasets.

Figure 3 shows the results of accuracybetween various datasets that includes PIMA Indian dataset, MDPP and CDI, where the results of DSVM achieves higher accuracy for classification of data samples than other methods.

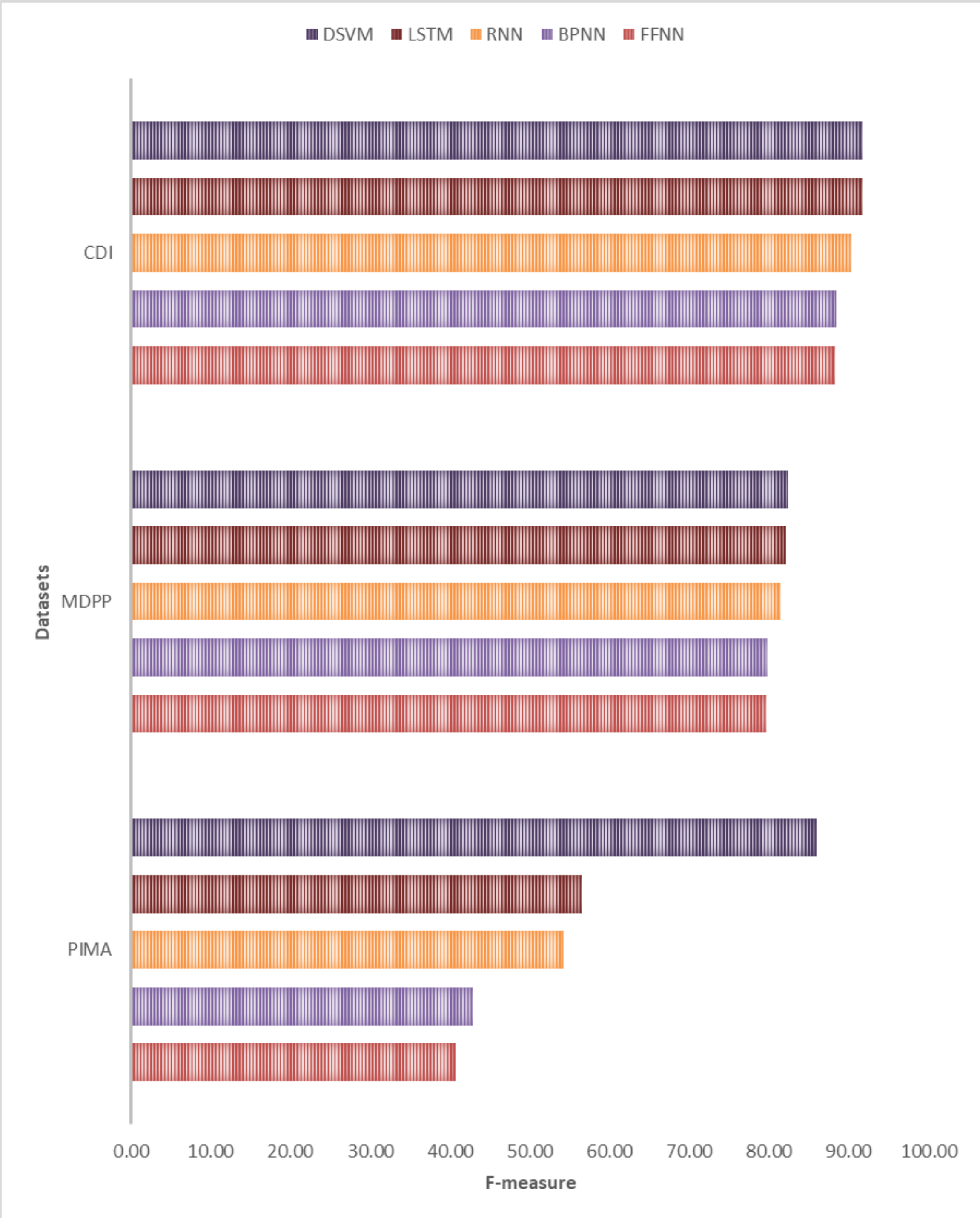

Figure 4: F-measure with various datasets.

Figure 4 shows the results of F-measure between various datasets that includes PIMA Indian dataset, MDPP and CDI, where the results of DSVM achieves higher F-measurefor classification of data samples than other methods.

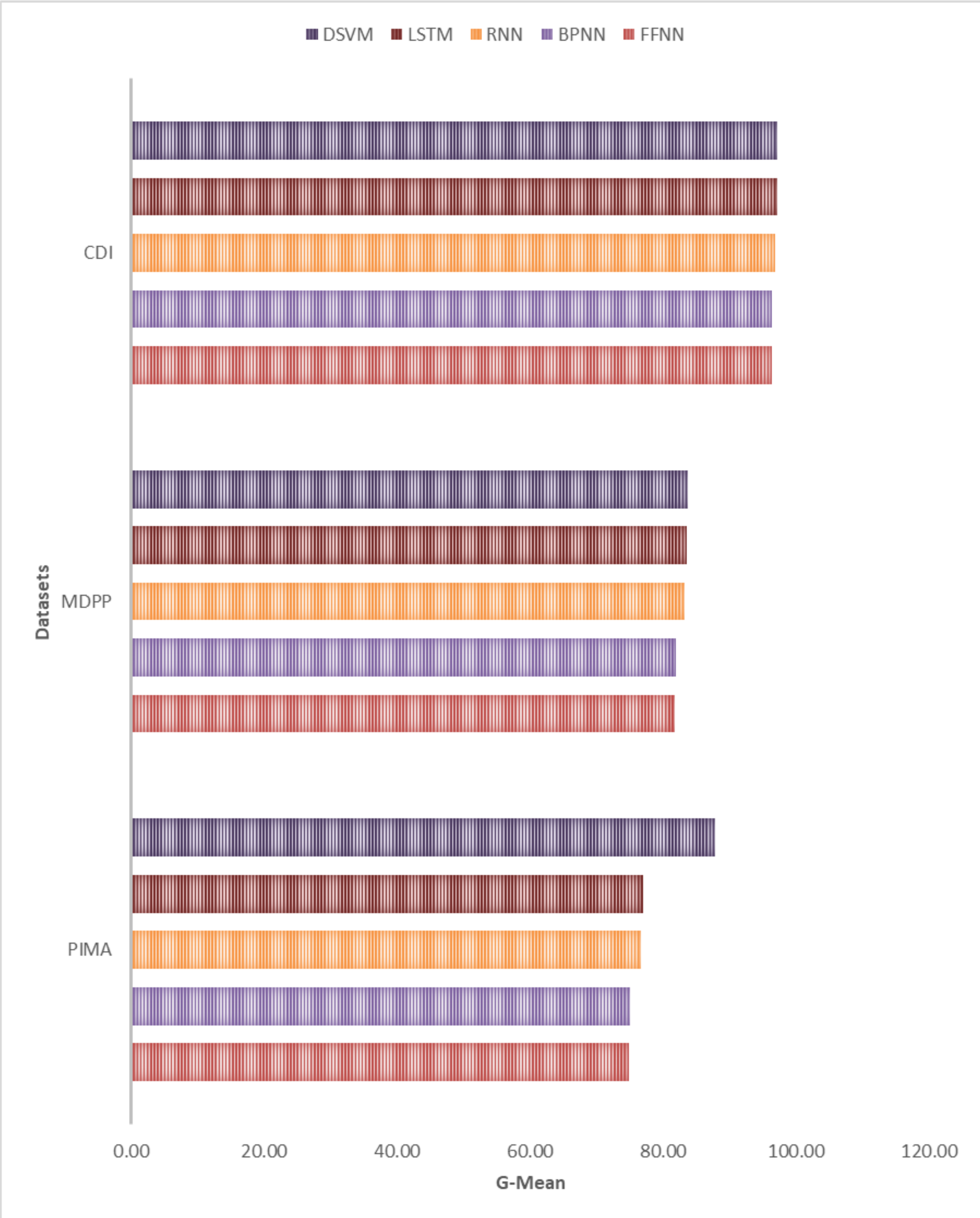

Figure 5: G-mean with various datasets.

Figure 5 shows the results of G-mean between various datasets that includes PIMA Indian dataset, MDPP and CDI, where the results of DSVM achieves higher G-meanfor classification of data samples than other methods.

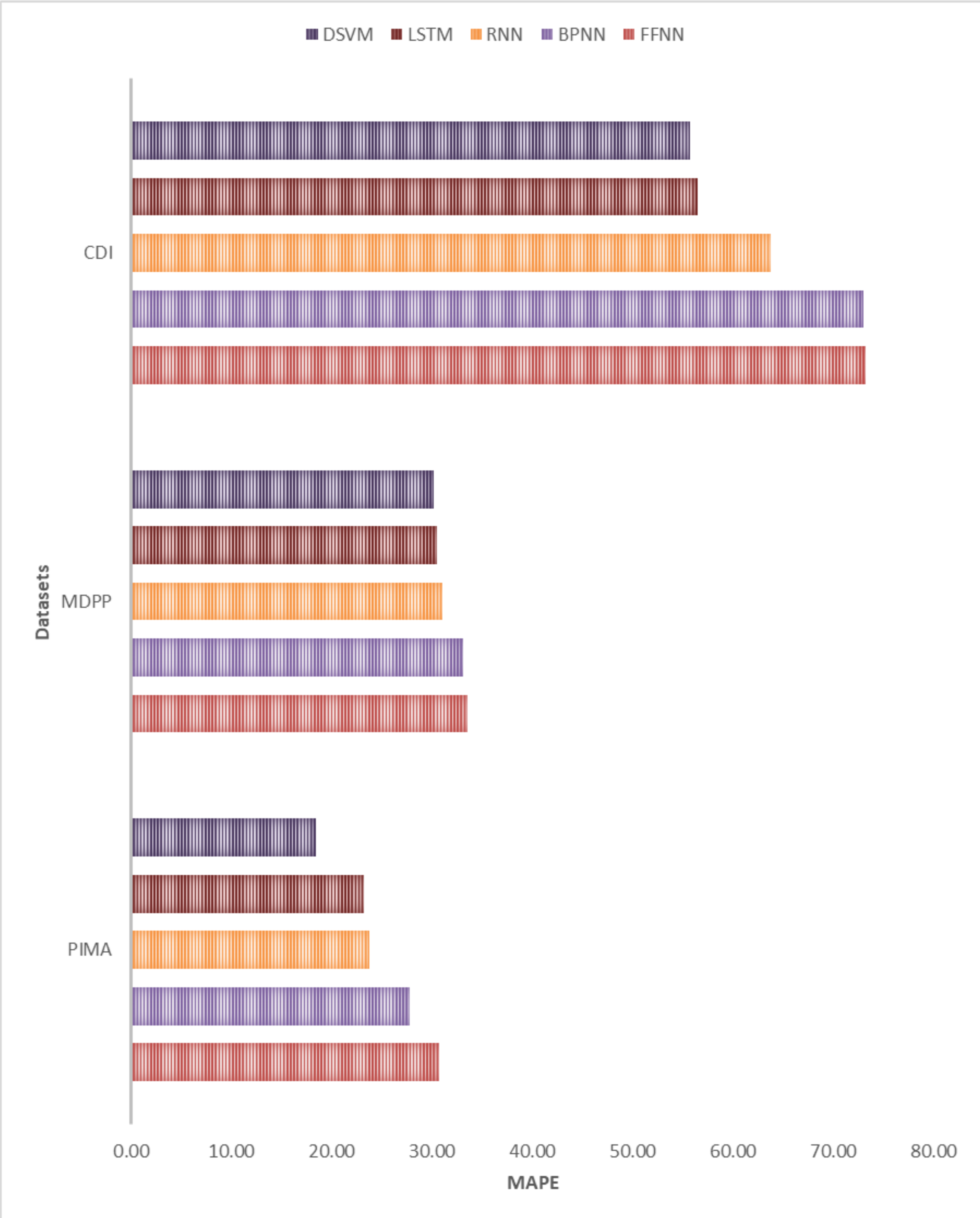

Figure 6: MAPE with various datasets.

Figure 6 shows the results of MAPE between various datasets that includes PIMA Indian dataset, MDPP and CDI, where the results of DSVM achieves reduced MAPE for classification of data samples than other methods.

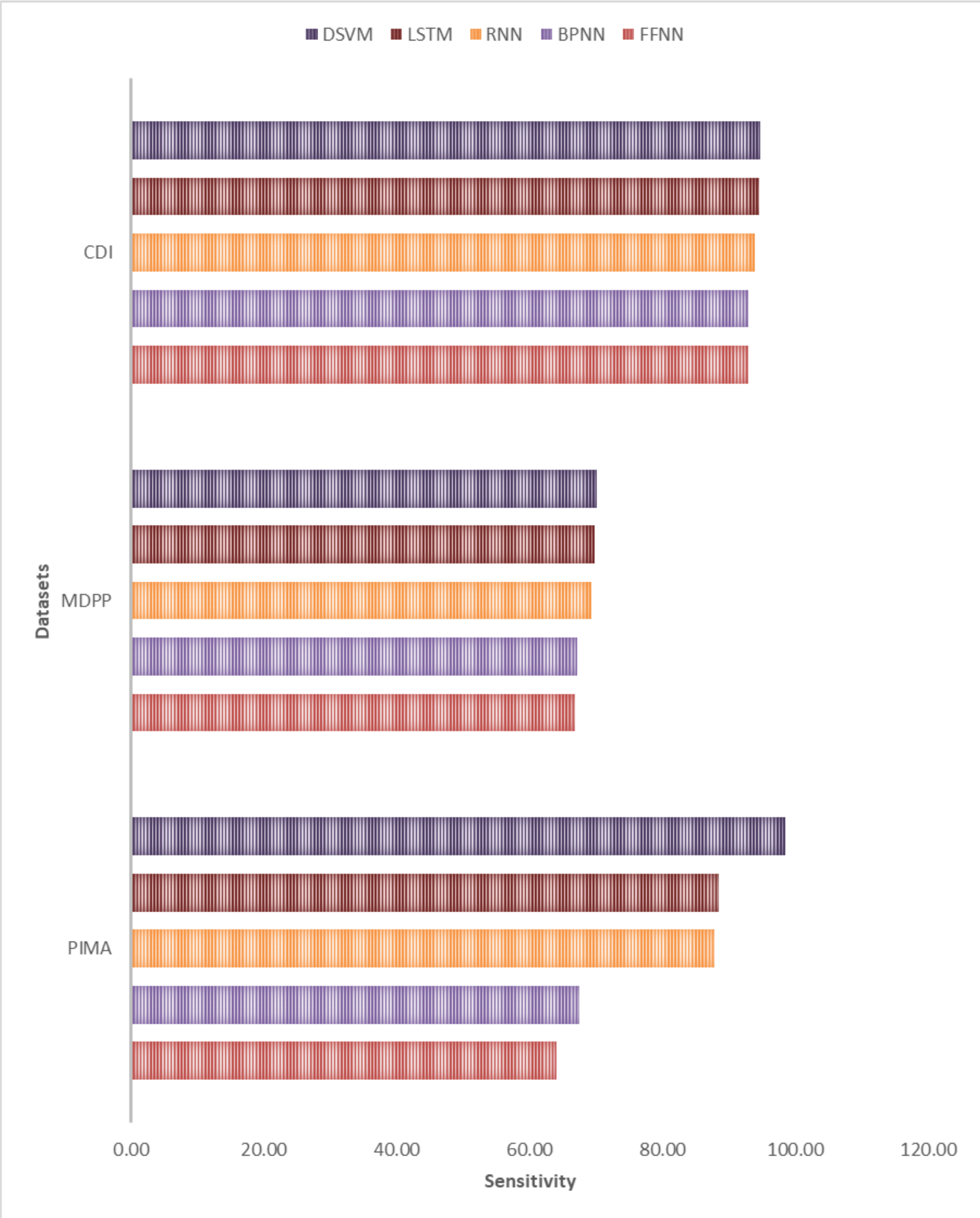

Figure 7: Sensitivity with various datasets.

Figure 7 shows the results of sensitivity between various datasets that includes PIMA Indian dataset, MDPP and CDI, where the results of DSVM achieves higher sensitivity for classification of data samples than other methods.

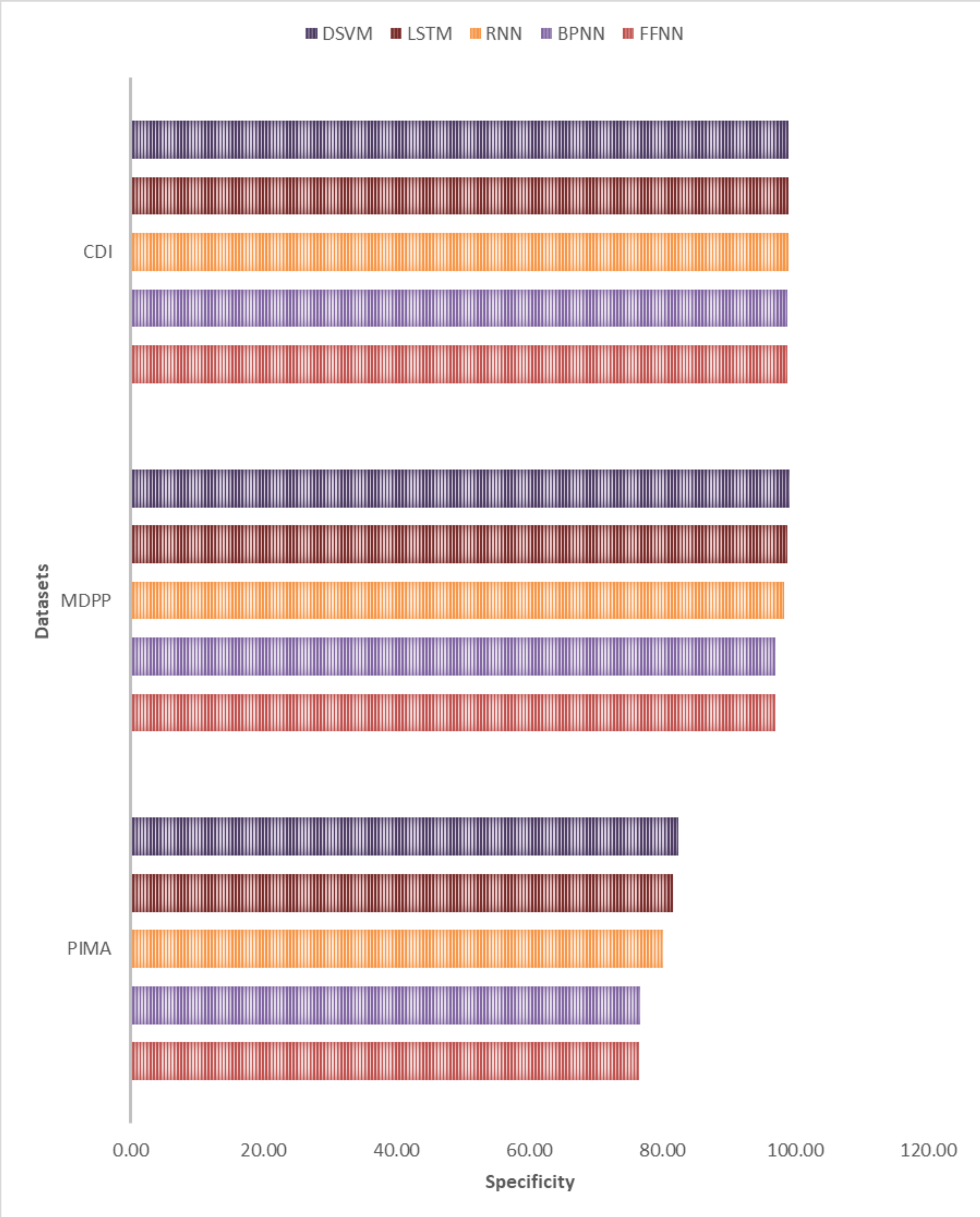

Figure 8: Specificity with various datasets.

Figure 8 shows the results of specificity between various datasets that includes PIMA Indian dataset, MDPP and CDI, where the results of DSVM achieves higher specificity for classification of data samples than other methods.

5. CONCLUSIONS

In this paper, the DSVM for classification considers various parameter of human body to classify the levels of diabetes mellitus. DSVM is used to classify the features after the feature extraction, and selection of instances in predicting the levels of diabetes mellitus. The simulation is conducted to test the efficacy of the model, and the results show that the deep learning paradigm provides increased accuracy, specificity, sensitivity and reduced mean absolute percentage error than other deep learning models.

REFERENCES

1. Das, H., Naik, B., & Behera, H. S. (2018). Classification of diabetes mellitus disease (DMD): a data mining (DM) approach. In *Progress in computing, analytics and networking* (pp. 539-549). Springer, Singapore.
2. Petersmann, A., Nauck, M., Müller-Wieland, D., Kerner, W., Müller, U. A., Landgraf, R., ... & Heinemann, L. (2018). Definition, classification and diagnosis of diabetes mellitus. *Experimental and clinical endocrinology & diabetes*, *126*(07), 406-410.
3. Maniruzzaman, M., Kumar, N., Abedin, M. M., Islam, M. S., Suri, H. S., El-Baz, A. S., & Suri, J. S. (2017). Comparative approaches for classification of diabetes mellitus data: Machine learning paradigm. *Computer methods and programs in biomedicine*, *152*, 23-34.
4. Petersmann, A., Müller-Wieland, D., Müller, U. A., Landgraf, R., Nauck, M., Freckmann, G., ... & Schleicher, E. (2019). Definition, classification and diagnosis of diabetes mellitus. *Experimental and Clinical Endocrinology & Diabetes*, *127*(S 01), S1-S7.
5. Daniel, A., Kannan, B. B., Yuvaraj, N., &Kousik, N. V. (2021). Predicting Energy Demands Constructed on Ensemble of Classifiers. In Intelligent Computing and Applications (pp. 575-583). Springer, Singapore.
6. Mercaldo, F., Nardone, V., &Santone, A. (2017). Diabetes mellitus affected patients classification and diagnosis through machine learning techniques. *Procedia computer science*, *112*, 2519-2528.

7. Raja, R. A., &Kousik, N. V. (2021). Privacy Preservation Between Privacy and Utility Using ECC-based PSO Algorithm. In Intelligent Computing and Applications (pp. 567-573). Springer, Singapore.
8. Alharbi, A., &Alghahtani, M. (2019). Using genetic algorithm and ELM neural networks for feature extraction and classification of type 2-diabetes mellitus. *Applied Artificial Intelligence*, *33*(4), 311-328.
9. Karthick, S., Yuvaraj, N., Rajakumari, P. A., & Raja, R. A. (2021). Ensemble Similarity Clustering Frame work for Categorical Dataset Clustering Using Swarm Intelligence. In Intelligent Computing and Applications (pp. 549-557). Springer, Singapore.
10. Mack, L. R., &Tomich, P. G. (2017). Gestational diabetes: diagnosis, classification, and clinical care. *Obstetrics and Gynecology Clinics*, *44*(2), 207-217.
11. Natarajan, Y., Raja, R. A., Kousik, D. N., & Johri, P. (2020). Improved Energy Efficient Wireless Sensor Networks Using Multicast Particle Swarm Optimization. Available at SSRN 3555764.
12. Sohail, N., Jiadong, R., Uba, M., Irshad, M., & Khan, A. (2018). Classification and cost benefit Analysis of Diabetes mellitus Dominance. *Int. J. Comput. Sci. Netw. Secur*, *18*, 29-35.
13. Wang, X., Zhai, M., Ren, Z., Ren, H., Li, M., Quan, D., ... &Qiu, L. (2021). Exploratory study on classification of diabetes mellitus through a combined Random Forest Classifier. *BMC medical informatics and decision making*, *21*(1), 1-14.
14. Wen, D., Zhou, Y., Li, P., Zhang, P., Li, J., Wang, Y., ... & Xu, Y. (2020). Resting-state EEG signal classification of amnestic mild cognitive impairment with type 2 diabetes mellitus based on multispectral image and convolutional neural network. *Journal of neural engineering*, *17*(3), 036005.
15. Raja, R. A., Yuvaraj, N., &Kousik, N. V. (2021). Analyses on Artificial Intelligence Framework to Detect Crime Pattern. Intelligent Data Analytics for Terror Threat Prediction: Architectures, Methodologies, Techniques and Applications, 119-132.
16. Wright Jr, J. R., & McIntyre, L. (2020). Misread and mistaken: Étienne Lancereaux's enduring legacy in the classification of diabetes mellitus. *Journal of medical biography*, 0967772020914797.
17. Beck-Nielsen, H. (2020). Methodologies in characterization and classification of prediabetic states to non-insulin-dependent diabetes mellitus (NIDDM). In *Research Methodologies in Human Diabetes Part 1* (pp. 147-170). De Gruyter.

18. Beck-Nielsen, H. (2020). Methodologies in characterization and classification of prediabetic states to non-insulin-dependent diabetes mellitus (NIDDM). In *Research Methodologies in Human Diabetes Part 1* (pp. 147-170). De Gruyter.
19. Choubey, D. K., Kumar, P., Tripathi, S., & Kumar, S. (2020). Performance evaluation of classification methods with PCA and PSO for diabetes. *Network Modeling Analysis in Health Informatics and Bioinformatics*, *9*(1), 1-30.
20. Islam, M. M., Rahman, M. J., Roy, D. C., &Maniruzzaman, M. (2020). Automated detection and classification of diabetes disease based on Bangladesh demographic and health survey data, 2011 using machine learning approach. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, *14*(3), 217-219.
21. Prasad, K. S., Reddy, N. C. S., & Puneeth, B. N. (2020). A Framework for Diagnosing Kidney Disease in Diabetes Patients Using Classification Algorithms. *SN Computer Science*, *1*(2), 1-6.
22. Gayathri, S., Krishna, A. K., Gopi, V. P., &Palanisamy, P. (2020). Automated binary and multiclass classification of diabetic retinopathy using haralick and multiresolution features. *IEEE Access*, *8*, 57497-57504.
23. PekelÖzmen, E., &Özcan, T. (2020). Diagnosis of diabetes mellitus using artificial neural network and classification and regression tree optimized with genetic algorithm. *Journal of Forecasting*, *39*(4), 661-670.
24. Reddy, G. T., Bhattacharya, S., Ramakrishnan, S. S., Chowdhary, C. L., Hakak, S., Kaluri, R., & Reddy, M. P. K. (2020, February). An ensemble based machine learning model for diabetic retinopathy classification. In *2020 international conference on emerging trends in information technology and engineering (ic-ETITE)* (pp. 1-6). IEEE.
25. Vargas, P., Moreles, M. A., Pena, J., Monroy, A., &Alavez, S. (2020). Estimation and SVM classification of glucose-insulin model parameters from OGTT data: a comparison with the ADA criteria. *International Journal of Diabetes in Developing Countries*, 1-9.