International Journal of Life Sciences, Biotechnology and Pharma Research Vol. 14, No. 4, April 2025 Online ISSN: 2250-3137
Print ISSN: 2977-0122

DOI: 10.69605/ijlbpr_14.4.2025.101

ORIGINAL RESEARCH

Prediction of Diabetes Mellitus Using Radial Basis Function, Recurrent Neural Network and Extreme Learning Machine algorithms

Mr. Jagdish D. Powar[1], Dr. Rajesh Dase[2], Dr. Deepak Bhosle[3]

[1]PhD student in Biostatistics, Department of Community Medicine, MGM, Medical College & Hospital, Chhatrapati Sambhaji Nagar, Maharashtra, India
[2]Assistant Professor (Statistics), Department of Statistics, SBES Science College Chhatrapati Sambhajinagar, Maharashtra, India
[3]Professor and Head, Department of Pharmacology, MGM, Medical College & Hospital, Chhatrapati Sambhaji Nagar, Maharashtra, India

Corresponding author
Mr. Jagdish D. Powar
PhD student in Biostatistics, Department of Community Medicine, MGM, Medical College & Hospital, Chhatrapati Sambhaji Nagar, Maharashtra, India
Email: jdpstat1479@gmail.com

Received: 11 March, 2025 Accepted: 30 March, 2025 Published: 10 April, 2025

ABSTRACT

Background: Diabetes mellitus (DM) is a chronic metabolic disorder with an increasing global prevalence that requires early detection and effective predictive models. The use of artificial intelligence in disease diagnosis and management has significantly increased in recent years. Artificial neural network (ANN) models playing a crucial role in enhancing predictive accuracy and clinical decision-making. This study compares the performance of Radial Basis Function (RBF) networks, Recurrent Neural Networks (RNNs), and Extreme Learning Machines (ELMs) for diabetes prediction. **Methods:** A case-control study was conducted on 800 participants (400 diabetics and 400 non-diabetics) recruited from a hospital outpatient department. The data were collected on demographic, lifestyle, anthropometric, and medical history variables using structured case record sheet.Predictive models were developed in RStudio and trained on 80% of the dataset, with 20% reserved for testing. Model performance was assessed using accuracy, sensitivity, specificity, Kappa score, and Area Under the Receiver Operating Characteristic (AUROC) curve. **Results:**In the prediction of diabetes, ELM outperformed RBF and RNN models in terms of accuracy (80.00%), sensitivity (77.50%), specificity (82.50%), Kappa score (0.60), and precision (81.58 percent).RBF attained moderate accuracy (74.38%), whereas RNN showed lower predictive performance (70.62%), likely due to its suitability for sequential data rather than structured medical datasets. **Conclusion:** ELM demonstrated superior predictive capability over RBF and RNN in diabetes classification, making it the recommended model for risk assessment.

Keywords: Diabetes Prediction, Artificial Neural Networks, RBF, RNN, ELM

INTRODUCTION

Diabetes mellitus (DM) is a chronic metabolic disorder characterized by high blood glucose levels due to inadequate insulin production or poor insulin use. According to the World Health Organization (WHO), diabetes is a major cause of death and morbidity around the world and its prevalence has increased significantly in recent decades (1). The International Diabetes Federation (IDF) estimates that by 2045 the global population of diabetics will reach 783 million people, putting a significant burden on health systems (2). With the second largest diabetes population in the world, early detection and treatment of diabetes is essential to reduce long-term complications (3).

Online ISSN: 2250-3137
Print ISSN: 2977-0122
DOI: 10.69605/ijlbpr_14.4.2025.101

Artificial Intelligence (AI) and Machine Learning (ML) have emerged as powerful tools in the early detection and prediction of diabetes. Traditional diagnostic methods, such as fasting plasma glucose (FPG) and oral glucose tolerance tests (OGTT), are widely used but can be time-consuming and resource-intensive. Recent advances in deep learning have introduced models like the Radial Basis Function (RBF) network, Recurrent Neural Networks (RNNs), and Extreme Learning Machine (ELM), which can analyze complex patient data patterns and enhance predictive accuracy (4,5).

RNNs are useful for identifying patterns and dependencies in medical records because they are well-suited to processing sequential data, especially Long Short-Term Memory (LSTM) networks (6). The ELM model, meanwhile, has been acknowledged for its high effectiveness in medical classification tasks and provides quick learning capabilities (7). A comparative study of these models can yield important information about the best method for predicting diabetes, which will ultimately help with early diagnosis and prompt treatment (8).

This study aims to develop and evaluate the performance of RBF, RNN, and ELM models for diabetes prediction using real-world patient data.

METHODOLOGY

A case-control study was conducted on individuals diagnosed with diabetes and healthy controls recruited from a hospital's Medicine Outpatient Department (OPD), with the goal of identifying predictive factors and developing diabetes prediction models. A total of 800 participants were included, with 400 diagnosed with diabetes and 400 healthy people who had no history or diagnosis of the condition. The sample size was calculated using the formula $n = 4pq/e^2$, based on the diabetes prevalence rate of 11% in India and a precision level of 5%. The final required sample size was 391. A structured case record form was used to collect comprehensive data, including demographic information, lifestyle habits, anthropometric measurements, and medical history. Age and gender were considered for matching between cases and controls.

Table 1: Details input variables and preprocessing of the dataset

Variable	Type	Categories / Range	Coding / Treatment	Data Processing
Diabetes	Categorical	No, Yes	No = 0, Yes =1	Label Encoding
Gender	Categorical	Male, Female	Male = 1, Female = 2	Label Encoding
Diet	Categorical	Mixed, Veg	Mixed = 1, Veg = 2	Label Encoding
Physical Activity	Categorical	No exercise, Mild, Moderate, Vigorous	No exercise = 1, Mild = 2, Moderate = 3, Vigorous = 4	Ordinal Encoding
Smoking	Categorical	Absent, Present	Absent = 1, Present = 2	Label Encoding
Alcohol Consumption	Categorical	No, Yes	No = 1, Yes = 2	Label Encoding
History of Diabetes	Categorical	None, One parent, Both	None = 1, One parent = 2, Both = 3	Ordinal Encoding
History of CVD	Categorical	No, Yes	No = 1, Yes = 2	Label Encoding
History of Covid	Categorical	No, Yes	No = 1, Yes = 2	Label Encoding
History of Hypertension	Categorical	None, One parent, Both	None = 1, One parent = 2, Both = 3	Ordinal Encoding
Age in Years	Numeric	20-60 years	-	Normalization
BMI	Numeric	15-40	-	Normalization
Waist Circumference	Numeric	60-120 cm	-	Normalization
Neck Circumference	Numeric	30-50 cm	-	Normalization
Stress Score	Numeric	0-40	-	Normalization

The categorical variables were converted to integer format, while the quantitative variables were normalized. The presence or absence of diabetes was used as an outcome variable in this study. Three predictive models for diabetes were developed: RBF, RNN, and ELM, with R studio software used for coding and model development. The models were trained on 80% of the dataset, with the remaining 20% used to test their performance. The performance of these models was assessed using key metrics such as accuracy, sensitivity, specificity, and the area under the receiver operating characteristic (ROC) curve.

Radial Basis Function

The radial basis function (RBF) is a type of activation function commonly used in artificial neural networks, in particular in neural networks with a radial base function and in support vector machines. This function is a real-valued function whose output depends only on distance from the center of the field,

DOI: 10.69605/ijlbpr_14.4.2025.101

which makes it very efficient for pattern recognition, functional approximation, and classification tasks. The Gaussian function, defined as r = e-γr, where r is the Euclidean distance between the input and the center of the center, is the most commonly used RBF parameter. RBF allows networks to learn complicated nonlinear relationships while preserving parallelism.

Recurrent Neural Network (RNN)

A recurrent neural network is a type of artificial neural network designed for parallel processing of data, which makes it highly efficient for tasks such as time series prediction, natural language processing and speech recognition. RNNs, unlike traditional transmission networks, have a recurrent connection, which allows them to store information from previous time steps, and thus to capture dependencies and patterns in sequence data. The hidden status is a key part of an RNN that acts like memory and updates at every step, based on current input and previous hidden status. But standard RNNs suffer from problems like vanishing and exploding gradients, which make learning long-term dependencies difficult. Advanced architectures such as long-term memory (LSTM) and gate recurrent units (GRUs) have been introduced to address this problem, and include gate mechanisms to control information flows and improve performance.

Extreme Learning Machine (ELM)

The Extreme Learning Machine (ELM) is a single-layer feedforward neural network learning algorithm that is quick and effective.Unlike traditional neural networks that use gradient-based optimization techniques, ELM randomly assigns weights and biases to hidden neurons and determines output weights analytically using Moore-Penrose pseudo-inverse. This eliminates the need for iterative weight updates, significantly reducing training time while maintaining high accuracy. ELM is known for its universal approximation capability, making it effective for classification, regression, and feature selection tasks. It overcomes issues like slow convergence and local minima found in traditional neural networks, making it a preferred choice for real-time applications such as image recognition, bioinformatics, and medical diagnosis.

RESULTS

Table 1: Association of Demographic, Lifestyle, and Family History Factors with Diabetes

Variable	Category	No Diabetes		Diabetes		Chi-Square	p-value
		n	%	n	%		
Gender	Male	200	50.0%	200	50.0%	0	1
	Female	200	50.0%	200	50.0%		
Age Groups	20-30	62	15.5%	62	15.5%	0.00	1.00
	30-40	147	36.8%	147	36.8%		
	40-50	102	25.5%	102	25.5%		
	50-60	89	22.3%	89	22.3%		
Diet	Mixed	103	25.8%	18	4.5%	68.71	<0.01
	Vegetarian	297	74.3%	382	95.5%		
Smoking	No	349	87.3%	366	91.5%	3.37	> 0.05
	Yes	51	12.8%	34	8.5%		
Alcohol Consumption	No	332	83.0%	344	86.0%	1.15	> 0.05
	Yes	68	17.0%	56	14.0%		
History of Diabetes in Parents	None	312	78.0%	273	68.3%	11.36	< 0.01
	One Parent	78	19.5%	119	29.8%		
	Both Parents	10	2.5%	8	2.0%		
History of Cardiovascular Disease	No	391	97.8%	378	94.5%	4.83	< 0.05
	Yes	9	2.3%	22	5.5%		
History of Covid	No	372	93.0%	361	90.3%	1.63	> 0.05
	Yes	28	7.0%	39	9.8%		
History of Hypertension in Parents	None	275	68.8%	279	69.8%	12.72	< 0.01
	One Parent	100	25.0%	115	28.8%		
	Both Parents	25	6.3%	6	1.5%		

The table presents the association between various demographic, lifestyle, and family history factors with diabetes using chi-square tests. Since age and gender matching were performed in this study, no significant association was observed for these variables ($p = 1.00$). However, significant associations were found for diet ($p < 0.01$), history of diabetes in parents ($p < 0.01$), history of cardiovascular disease ($p < 0.05$), and history of hypertension in parents ($p < 0.01$), indicating their potential role in diabetes risk. A family history of diabetes was also a key predictor, with individuals having one diabetic parent showing a higher prevalence (29.8%) compared to those with no parental history (68.3%). Similarly, individuals with a

DOI: 10.69605/ijlbpr_14.4.2025.101

history of cardiovascular disease (5.5%) and hypertension in both parents (1.5%) showed significant associations with diabetes. However, smoking ($p > 0.05$), alcohol consumption ($p > 0.05$), and history of COVID-19 ($p > 0.05$) did not show statistically significant associations. These findings highlight the importance of dietary habits and family medical history in assessing diabetes risk.

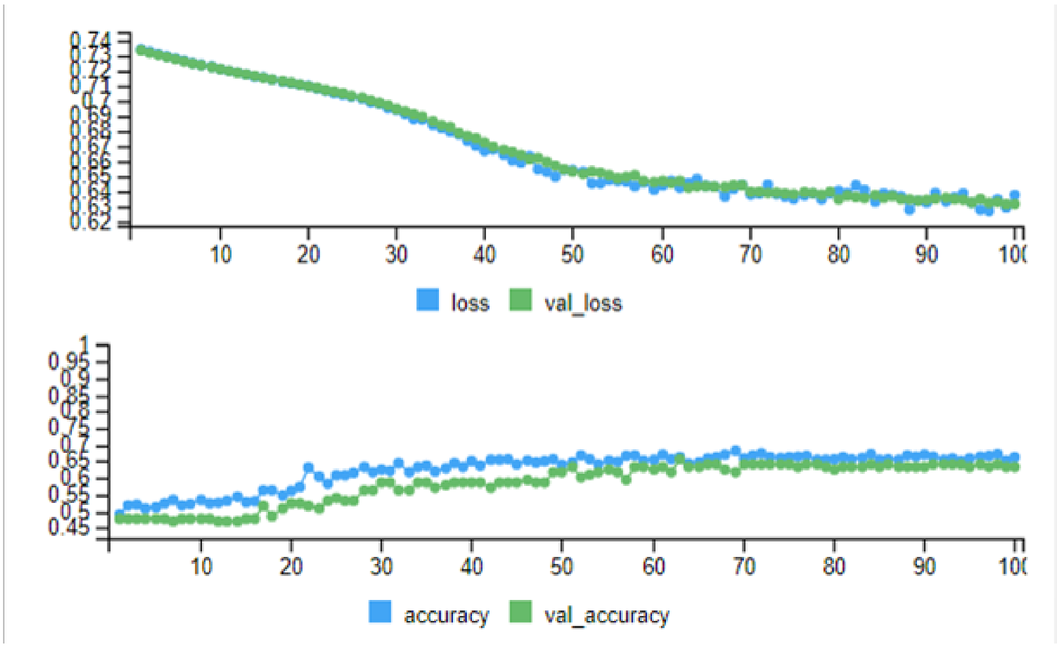

Figure 1: Training and validation performance of the RNN model

The plots show the training and validation performance of the RNN model over 100 epochs. The loss decreases steadily, indicating learning, while the accuracy gradually improves. The validation curves closely follow the training curves, suggesting minimal overfitting. However, fluctuations in validation accuracy indicate some variability, highlighting the need for further fine-tuning.

Table 3: Comparison of Performance Metrics for RBF, RNN, and ELM in Diabetes Prediction

Metric	RBF	RNN	ELM
Accuracy	74.38%	70.62%	80.00%
Kappa Score	0.49	0.41	0.6
Sensitivity (Recall)	70.00%	61.25%	77.50%
Specificity	78.75%	80.00%	82.50%
Positive Predictive Value (Precision)	76.71%	75.38%	81.58%
Negative Predictive Value	72.41%	67.37%	78.57%

The table compares the performance of three Artificial Neural Network models—RBF, RNN and ELMin predicting diabetes using various evaluation metrics. Among the models, ELM demonstrated the highest accuracy (80.00%), followed by RBF (74.38%) and RNN (70.62%). The Kappa score, which measures agreement beyond chance, was also highest for ELM (0.60), indicating better reliability compared to RBF (0.49) and RNN (0.41). In terms of sensitivity (recall), which reflects the model's ability to correctly identify diabetic patients, ELM achieved the highest score (77.50%), followed by RBF (70.00%) and RNN (61.25%). Specificity, which measures the ability to correctly identify non-diabetic individuals, was highest for ELM (82.50%), with RBF and RNN scoring 78.75% and 80.00%, respectively. ELM also outperformed the other models in positive predictive value (precision) at 81.58%, indicating fewer false positives compared to RBF (76.71%) and RNN (75.38%). Similarly, ELM had the highest negative predictive value (78.57%), meaning it was more effective in correctly predicting non-diabetic cases compared to RBF (72.41%) and RNN (67.37%). Overall, these results suggest that ELM provides the most robust performance in diabetes prediction, outperforming RBF and RNN across all key metrics.

DISCUSSION

Three neural network models—Radial Basis Function (RBF), Recurrent Neural Network (RNN), and Extreme Learning Machine (ELM)—were compared in this study for their ability to predict diabetes mellitus. The findings showed that ELM performed better than RBF and RNN, with the highest kappa score (0.6), sensitivity (77.50 percent), specificity (82.50 percent), accuracy (80.00 percent), positive

DOI: 10.69605/ijlbpr_14.4.2025.101

predictive value (81.58 percent), and negative predictive value (78.57 percent). These results are in line with other studies that show how beneficial ELM is for medical diagnostics. Numerous studies have assessed deep learning and machine learning models for predicting diabetes. Joshi (9) et.al. reported that neural network-based methods outperformed traditional statistical models in terms of predictive performance after implementing a variety of machine learning techniques. The results of their study showed that ELM was effective in predicting diabetes, with an accuracy of 82.1 percent. Similar to this, an early-stage diabetes prediction study that used ELM highlighted its quick training time and strong generalization ability by reporting an accuracy of 85.3 percent (10)

The effectiveness of RBF networks has been explored in diabetes classification as well. Venkatesan and Anitha(11) applied an RBF neural network and achieved an accuracy of 76.8%, which aligns closely with the 74.38% accuracy observed in our study. While RBF networks demonstrated reasonable predictive capability, they were outperformed by ELM. This is in agreement with studies suggesting that RBF networks may struggle with large-scale datasets due to their localized kernel-based approach.

Recurrent Neural Networks (RNNs) have been applied in various medical predictions, including diabetes-related complications. A study evaluating RNNs for predicting cardiovascular disease in diabetic patients achieved an accuracy of 73.5% (12), which is slightly higher than the 70.62% accuracy observed in our study. RNNs, although effective in processing sequential data, may not be the best choice for structured medical datasets where feature interactions are not necessarily sequential.

Several ensemble learning techniques have been explored to enhance diabetes prediction accuracy. A study utilizing boosting algorithms on the Pima Indian Diabetes Dataset reported improved accuracy of up to 79.5% (13), suggesting that ensemble methods could provide benefits similar to those observed with ELM. Another study integrating particle swarm optimization with neural networks (PSO-NNDP) achieved an accuracy of 83.7% (14), reinforcing the importance of hybrid models in improving predictive performance.

Additionally, research integrating artificial intelligence with diagnostic tools like Raman spectroscopy has shown promising results. One study demonstrated that combining machine learning with Raman spectroscopy could achieve an accuracy of 86.4% in diabetes screening (15). Our results are consistent with previous research showing ELM's superior predictive ability in diabetes diagnosis. ELM is the recommended option for diabetes risk assessment due to its capacity to generalize across various scenarios, even though RBF and RNN models did fairly well. To further improve predictive accuracy, future studies should investigate hybrid models that combine ensemble techniques or optimization algorithms with ELM. Furthermore, combining AI with biochemical profiling or diagnostic imaging could improve early diabetes detection techniques even more.

CONCLUSION

According to this study, Extreme Learning Machines performed better than Radial Basis Function and Recurrent Neural Network models in predicting diabetes mellitus. ELM achieved the highest kappa score (0.6), sensitivity (77.50%), specificity (82.50%), and accuracy (80.00%). For diabetes risk assessment, ELM is the recommended model; however, in order to increase prediction accuracy and early diagnosis, future research should concentrate on hybrid models and multimodal AI techniques.

REFERENCES

1. World Health Organization. Diabetes. Available from: https://www.who.int/news-room/fact-sheets/detail/diabetes.
2. International Diabetes Federation. IDF Diabetes Atlas, 10th edition. 2021. Available from: https://diabetesatlas.org.
3. Indian Council of Medical Research. Guidelines for Management of Type 2 Diabetes. 2022. Available from: https://www.icmr.nic.in.
4. Hinton G, Deng L, Yu D, et al. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. IEEE Signal Process Mag. 2012;29(6):82-97.
5. Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput. 1997;9(8):1735-1780.
6. Huang GB, Zhou H, Ding X, Zhang R. Extreme learning machine for regression and multiclass classification. IEEE Trans Syst Man Cybern B Cybern. 2012;42(2):513-529.
7. Kavakiotis I, Tsave O, Salifoglou A, Maglaveras N, Vlahavas I, Chouvarda I. Machine learning and data mining methods in diabetes research. Comput Struct Biotechnol J. 2017;15:104-116.
8. Shatte AB, Holdsworth J, Lee I. Machine learning in mental health: a scoping review of methods and applications. Psychol Med. 2019;49(9):1426-1448.
9. Joshi TN, et al. Diabetes Prediction Using Machine Learning Techniques. *J Cardiovasc Dis Res*. 2023;14(2):91-97. Available from: https://jcdronline.org/admin/Uploads/Files/65f5a0890a0220.41907535.pdf
10. Elsayed N, ElSayed Z, Ozer M. Early Stage Diabetes Prediction via Extreme Learning Machine. *arXiv preprint*. 2022 Feb 22;arXiv:2202.11216. Available from: https://arxiv.org/pdf/2202.11216
11. Venkatesan P, Anitha S. Application of a Radial Basis Function Neural Network for Diagnosis of Diabetes Mellitus. *Curr Sci*. 2006;91(9):1195-1200. Available from: https://core.ac.uk/download/pdf/83048963.pdf
12. Kim J, et al. Prediction of Cardiovascular Complications in Patients with Newly Diagnosed Type 2 Diabetes Mellitus Using Machine Learning-Based Risk Engine. *Endocrinol Metab (Seoul)*. 2023;38(1):54-63. Available from: https://www.e-enm.org/journal/view.php?doi=10.3803%2FEnM.2023.1739

Online ISSN: 2250-3137
Print ISSN: 2977-0122

DOI: 10.69605/ijlbpr_14.4.2025.101

13. Gupta S, et al. An Ensemble Learning Approach for Diabetes Prediction Using Boosting Algorithms. *BMC Med Inform DecisMak*. 2023;23(1):251. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC10639159/
14. Kumar R, et al. Bio-Inspired PSO for Improving Neural-Based Diabetes Prediction System. *J ICT Standardization*. 2023;11(1):53-70. Available from: https://journals.riverpublishers.com/index.php/JICTS/article/view/12437
15. Omar M, et al. Use of Raman Spectroscopy to Screen Diabetes Mellitus with Machine Learning-Based Analysis Algorithms. *PLoS One*. 2018;13(10): e0205867. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC6179393/