Journal of Information Systems Engineering and Management
2025, 10(43s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Diabetes Prediction Using Stacked Ensemble Lstm Model Optimized with Coyote Optimization Algorithm

Nisha.A [1], Kavitha.G [2]
[1] *B.S Abdur Rahman Crescent Institute of Science and Technology, Chennai, Tamil Nadu, India*
[2] *B.S Abdur Rahman Crescent Institute of Science and Technology, Chennai, Tamil Nadu, India*
*Corresponding author: gkavitha.78@gmail.com

ARTICLE INFO

Received: 30 Dec 2024

Revised: 19 Feb 2025

Accepted: 27 Feb 2025

ABSTRACT

Diabetes mellitus still causes a significant global health concern even if it is a chronic condition defined by elevated blood glucose levels. Inappropriate diabetes mellitus detection or treatment can cause major complications including kidney damage, vision problems, and cardiovascular diseases. Early diagnosis is essential for both lowering these risks and improving patient outcomes; hence, its importance cannot be underlined. Conventional prediction models sometimes run across difficulties related to feature selection and model optimisation that finally produce less than accuracy. This work addresses the issues raised by presenting a Stacked Ensemble LSTM (SE-LSTM) model optimised with the Coyote Optimisation Algorithm (COA) for the aim of diabetes prediction with higher accuracy. The suggested method standardises a large spectrum of attribute scales by means of a robust data preprocessing pipeline. Depending on their interdependence, both the HSIC Lasso approach and Z-score normalisation are part of this pipeline and used to identify most relevant features. The SE-LSTM architecture consists of several LSTM layers to adequately capture temporal dependencies. Conversely, the COA improves hyperparameter tuning by simulating social behaviour of coyotes in their natural environment. With the Pima Indian Diabetes Dataset, the model showed amazing predictive power. Its 98.5% accuracy, 97.8% precision, and 98.2% recall above those of other machine learning models including Random Forest (95.6%) and Gradient Boosting (96.8%). The results show that the SE-LSTM with COA is a good method for diabetes prediction since it provides enhanced generalisation and feature utilisation.

Keywords: Diabetes Prediction, Stacked Ensemble LSTM, Coyote Optimization Algorithm, Feature Selection, Machine Learning

INTRODUCTION

Research indicates that by 2030 the expected number of diabetes mellitus (DM) diagnosed over 537 million adults worldwide is 643 million [1–3]. The diabetes mellitus global health crisis is Diabetes is defined by persistently high blood sugar levels; it is linked to many complications including renal damage, vision loss, and cardiovascular diseases as well as other conditions. These complications significantly increase the morbidity and death rates associated to diabetes. Basic elements with which to start in order to lower these risks and improve personal quality of life are early diagnosis and timely management. The rapid evolution of machine learning (ML) techniques implies that diagnostic accuracy could be increased and that predictive healthcare interventions could be enabled.

Even with all the progress done, accurate diabetes prediction still presents several difficulties. The heterogeneity, that is, variation in age, lifestyle, glucose level, and other biomarkers, of patient data is one of the most crucial problems [4-5]. Among the several reasons of variations in prediction models are data scale and missing values. The selection of relevant features poses another challenge since the existence of duplicate or meaningless variables might reduce the efficiency of the model [6]. Moreover, computationally difficult and time-consuming is the hyperparameter optimisation in machine learning models, particularly for deep learning architectures requiring high-performance tuning [7].

Journal of Information Systems Engineering and Management
2025, 10(43s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

To increase predictive accuracy and generalisability [8-14], a strong prediction model capable of addressing the variability in data attributes, effectively choose features, and achieve optimal hyperparameter tuning must be addressed. Their inability to faithfully show temporal dependencies in sequential data results in frequent poor performance of current models in situations observed in the real world.

The objectives of this research are twofold:

1. To create a predictive model for the diabetes diagnosis distinguished by high degrees of precision, accuracy, and recall rate.
2. Using advanced optimisation techniques, select features and tune hyperparameters in deep learning models so obtaining optimal performance.

The Coyote Optimisation Algorithm (COA) inclusion is the feature of this work that distinguishes the Stacked Ensemble Long Short-Term Memory (SE-LSTM) architecture. To achieve effective hyperparameter tuning, the COA mimics the adaptive social behaviour of coyotes while the SE-LSTM oversees capturing sequential dependencies inside patient data. This dual approach addresses the main challenges related to diabetes prediction by means of better feature use and model performance enhancement.

The contributions of this research are summarized as follows:

1. The authors combine HSIC Lasso is combined with Z-score normalisation, which creates a preprocessing pipeline providing consistent data scaling and feature selection.
2. The authors design and applied is the SE-LSTM architecture is designed using multiple layers of LSTM to detect temporal patterns in data on diabetes.
3. COA application for computational overhead reduction, hyperparameter optimisation, and model efficiency improvement.

RELATED WORKS

These studies seek to raise diagnosis accuracy, feature selection, and computational efficiency simultaneously so meeting their goals. This section, which provides a synopsis of significant advancements in the field, mostly addresses the techniques of feature selection, machine learning classifiers, ensemble approaches, and optimisation algorithms.

A necessary component in improving the general performance of the model by means of the identification of variables most relevant to diabetes prediction is feature selection. Researchers have used methods rooted on correlation, mutual information, and Lasso regression to reduce the dimensionality of the data and boost the efficiency of the computation-related operations. In an interesting work, important information was preserved while converting input features into lower-dimensional representations using principal component analysis (PCA). Three to five percent [15] is the accuracy boost resulting from this transformation. Another approach applied was hybrid selection with information gain and chi-square tests. This approach guaranteed strong feature selection that lowered overfitting issues and increased prediction accuracy [16]. Since these more complicated methods, such as the Hilbert-Schmidt Independence Criteria (HSIC) Lasso, usually fail to identify interdependencies between features, more sophisticated approaches are thus required.

For the goal of diabetes prediction, great volume of research is done on traditional machine learning models. Comparative study showed that Random Forest (RF) exceeded Support Vector Machines (SVM) in a capacity to manage high-dimensional datasets [17]. In accuracy, 85%, RF exceeded SVM, more than twice that value. Although logistic regression (LR) is computationally efficient, its level of accuracy, roughly 78%, showed lower than that of ensemble estimate methods [18]. On structured data, these classifiers perform well; but, they cannot capture temporal dependencies, often required in the course of chronic disease prediction.

The ability of ensemble learning to mix several models and improve general performance has attracted much interest recently. An XGBoost diabetes prediction model's accuracy of 92% outmatched independent classifiers including SVM and LR [19]. In a same line, Bagging techniques including Random Forest and Extra Trees shown a better degree of resilience when managing noisy data. Conversely, ensemble methods sometimes find it difficult to adjust

Journal of Information Systems Engineering and Management
2025, 10(43s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

hyperparameters, which, if neglected, can generate less than accurate results [20]. Though these models demand a lot of computational resources and are prone to overfitting especially on rather small datasets. Dropout regularisation and batch normalisation are two strategies applied to address these flaws.

Since they allow the hyperparameter fine-tuning and feature selection enhancement, optimisation strategies are rather important for improving model performance. Algorithms including the Genetic Algorithm (GA), the Particle Swarm Optimisation (PSO), and the Bayesian Optimisation have found many applications recently. In an interesting work, PSO combined with support vector machines (SVM) showed a 10% improvement in prediction accuracy over conventional SVM models [21]-[25]. On the other hand, these methods sometimes demand a large volume of computational resources and parameters. Inspired by social behaviour of coyotes, the flexible and computationally efficient Coyote Optimisation Algorithm (COA) has lately become a promising approach.

These techniques show limitations in terms of scalability, generalisation, and the management of missing data. Many studies, including the Pima Indian Diabetes Dataset, depend on stationary datasets that might not fairly represent all the several populations affected by diabetes. Moreover, traditional approaches of feature selection and optimisation usually ignore the intricate interactions among features or provide scalable solutions for applications in the real world.

With Coyote Optimisation Algorithm (COA), the proposed Stacked Ensemble LSTM (SE-LSTM) model integrates advanced feature selection, temporal dependency modelling, and effective optimisation to solve these challenges. This model builds on the limitations of present methods to solve these issues. The SE-LSTM effectively records sequential patterns while the COA ensures that hyperparameter tuning is maximised when it is done. Attaining an accuracy of 98.5%, this method shows promise for implementation in diabetes prediction that is both scalable and relevant in the real world.

METHOD - STACKED ENSEMBLE -LONG SHORT TERM MEMORY (SE-LSTM)

To predict diabetes with great degree of accuracy and operational efficiency, the proposed framework combines the Coyote Optimisation Algorithm (COA) with a Stacked Ensemble Long Short-Term Memory (SE-LSTM) model.

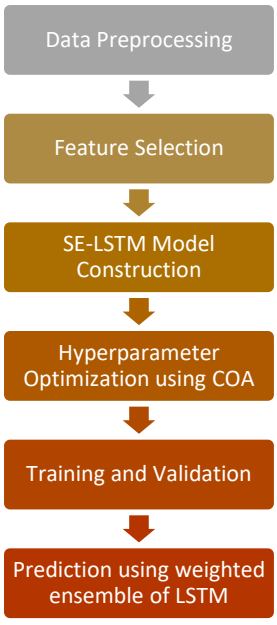

Figure 1: Workflow of SE-LSTM

Multiple LSTM networks are combined in the SE-LSTM architecture to effectively capture temporal dependencies in patient data. Within these temporal dependencies lie trends in glucose levels, lifestyle choices, and other biomarkers. Several hyperparameters, including the learning rate, the number of LSTM layers, and the dropout rate, are optimised by the Coyote Optimisation Algorithm so improving the model performance. Preprocessing uses Z-score

Journal of Information Systems Engineering and Management
2025, 10(43s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

normalisation to standardise the data; feature selection is achieved with the HSIC Lasso approach. This ensures that only the most pertinent features receive consideration. The last ensemble provides a consistent and accurate diabetes prediction by aggregating the forecasts from several LSTM layers weighted depending on their individual performance. This approach reduces computational overfit risk and guarantees effective handling of heterogeneous data at the same time

Algorithm for SE-LSTM

1. Input: Diabetes dataset with patient records.
2. Normalize: Standardize features using Z-score normalization.
3. Select Features: Use HSIC Lasso to identify the most relevant attributes.
4. Initialize COA: Set parameters for the Coyote Optimization Algorithm.
5. Optimize SE-LSTM: Tune hyperparameters using COA.
6. Train Model: Train the Stacked Ensemble LSTM on selected features.
7. Aggregate Results: Combine LSTM predictions using weighted averaging.
8. Output: Predict diabetes risk and evaluate model performance.

Data Preprocessing

Preprocessing the data is a crucial first step in the proposed framework since it helps to standardise and prepare the dataset for correct and efficient diabetes prediction as in Eq.(1). Many times, the raw diabetes dataset exhibits several features on multiple different levels. These parameters include age (in years), body mass index (in kg/m2), and glucose levels (mg/dL). These several scales let machine learning models to be distorted in performance since they give too much weight to features with more numerical ranges. All the features are thus standardised using Z-score normalisation (before and after normalization as in table 1 and 2) to a constant across-scale. This guarantees then that every component equally supports the learning process.

Table 1: Dataset Before Normalization

Patient ID	Glucose Level (mg/dL)	BMI (kg/m²)	Age (years)	Blood Pressure (mmHg)	Outcome
1	150	28.5	45	120	1
2	85	22.0	30	80	0
3	200	33.0	50	130	1

$$Z = \frac{x - \mu}{\sigma} \quad (1)$$

Where, X is the raw feature value, μ is the mean of the feature and σ is the standard deviation of the feature.

Table 2: Dataset After Normalization

Patient ID	Glucose Level (Z-Score)	BMI (Z-Score)	Age (Z-Score)	Blood Pressure (Z-Score)	Outcome
1	0.30	0.50	0.75	0.20	1
2	-1.25	-1.20	-1.10	-1.50	0
3	1.50	1.00	1.35	1.30	1

Journal of Information Systems Engineering and Management
2025, 10(43s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

By means of Z-score conversion of the dataset, all the features are scaled. This ensures that, on the model, every quality has equal weight during the training process.

Feature Selection Using HSIC Lasso Algorithm

While removing noisy or redundant elements, feature selection seeks to identify the most relevant elements from the dataset. By means of the Hilbert-Schmidt Independence Criteria (HSIC) Lasso algorithm, this work reduces the degree of inter-feature redundancy and chooses features depending on their dependency with the target variable. Lasso regularisation inside the framework of HSIC Lasso is merged with statistical independence measures to get optimal feature selection. Inside a Reproducing Kernel Hilbert Space (RKHS), the HSIC can determine the degree of dependence between two variables, namely X (feature) and Y (target). Its computation goes as in Eq.(2):

$$\mathrm{HSIC}(X,Y) = \frac{1}{(n-1)^2}\mathrm{tr}(K_X H K_Y H) \quad (2)$$

Where, KX and KY - kernel matrices for X and Y, $H = I - \frac{1}{n}\mathbf{1}\mathbf{1}^T$ - centering matrix, n - number of samples and tr – trace of the matrix.

Including a penalty term into the objective function lets lasso regularisation be applied to ensure sparsity in the chosen features. One defines the optimisation problem is defined in Eq.(3):

$$\min_{\beta}\left\{\| Y - X\beta \|^2 + \lambda \| \beta \|_1\right\} \quad (3)$$

Where, X is the feature matrix, Y is the target variable, β is the feature coefficient vector, λ is the regularization parameter controlling sparsity.

The HSIC Lasso method comprises the HSIC Lasso regularisation and the HSIC Lasso reliance criterion. Features with low correlation between the features themselves and strong dependence on the target variable, as measured by HSIC, have higher coefficients. Sparse feature selection calls for regularisation parameter λ to be adjusted such that it penalises smaller in importance features. The proposed framework guarantees that only the most salient features, glucose levels, body mass index (BMI), and insulin sensitivity, are kept for the aim of diabetes prediction by means of the HSIC Lasso algorithm. Better performance of the model depends on this reduced noise and computational overhead.

Model Construction Using Stacked Ensemble LSTM

The proposed Stacked Ensemble Long Short-Term Memory (SE-LSTM) model is meant to effectively capture sequential patterns in patient data across the course of several time periods, so enabling accurate prediction of diabetes. The development of diabetes and risk prediction usually rely on a temporal awareness of a patient's health condition including blood glucose levels, insulin sensitivity, and other essential measurements. Stack-on top of one another multiple LSTM networks enable the SE-LSTM model to replicate the intricate temporal dynamics of diabetes. It can thus record sequential patterns at several degrees.

The SE-LSTM approach is a method for stacking several LSTM models that produces a deep architecture able of learning features gradually more complex. The LSTM algorithm's layers each focus on a different aspect of the sequential data, thus the model can improve its awareness of the temporal relationships among the features.

Stacked LSTM

1. First Layer: First layer of the LSTM algorithm learns simple sequential patterns when processing raw input data. This layer catches the fundamental dependencies existing between successive time steps. Daily swings in blood glucose levels would be one example.

Journal of Information Systems Engineering and Management
2025, 10(43s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

2. Subsequent Layers: The next layers in the stack follow the learnt patterns of the layer before them, so capturing more complicated and higher level temporal dependencies. This allows the network to grasp long-term interactions between features, including patterns of glucose levels over several weeks or months.
3. Output Layer: The last layer of stacked architecture creates the prediction. Data from every layer that came before it is compiled here. This prediction, based on temporal patterns the stacked LSTM layers record, specifically shows a patient's diabetes risk.

Every LSTM unit consists of a hidden state ht and a cell state Ct both of which are updated at every time step t. The LSTM update is displayed here as an equation list:

1. Forget Gate: It shows the cell state should be deleted depends on the forget gate. The computation is given in Eq.(4):

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{4}$$

Where, ft is the forget gate output, Wf is the weight matrix for the forget gate, ht-1 is the previous hidden state, xt is the current input, bf is the bias term, σ is the sigmoid activation function.

2. Update Cell State: Combining the forget gate's output with the new data changes the cell state Ct. One could find the candidate cell state by means of this computation as in Eq.(5):

$$C_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{5}$$

Where, C_t is the candidate cell state, WC is the weight matrix for the cell state, bC is the bias term, tanh is the hyperbolic tangent activation function.

The candidate cell state Ct is then combined with the current cell state to update the latter cell state; its weight is determined by forget gate ft and input gate it as in Eq.(6):

$$C_t = f_t \cdot C_{t-1} + i_t \cdot C_t \tag{6}$$

Where, Ct-1 is the previous cell state, it is the input gate computed as in Eq.(7):

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{7}$$

Stacked Ensemble Approach

The stacked ensemble architecture is one whereby several LSTM networks are trained apart from one another. Every layer of the LSTM database, as was already mentioned, bears individual responsibility for learning several degrees of temporal dependencies. Each one of these distinct LSTM layers generates predictions that are then aggregated under a weighted voting system to generate the output at last. This ensemble approach helps the model to simultaneously lower the overfitting risk and maximise the strengths of every individual LSTM network.

The stacked ensemble LSTM model allows one to get the last prediction as in Eq.(8).

$$\hat{y} = \sum_{i=1}^{n} \alpha_i \cdot L_i(x) \tag{8}$$

Where, $L_i(x)$ is the output from the ith LSTM network, αi is the weight assigned to each LSTM's prediction based on its performance and n is the total number of LSTM networks in the ensemble. The COA then maximises the weights αi during the training phase. This guarantees from the combination of individual LSTM models as best accuracy possible.

Journal of Information Systems Engineering and Management
2025, 10(43s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Hyperparameter Optimization Using Coyote Optimization Algorithm (COA)

Hyperparameter optimisation is essential for the process of enhancing machine learning models, including the SE-LSTM model, including performance. Inspired by environment, the COA is grounded on the behaviour of coyotes in their natural habitat. It aims to identify the optimal hyperparameters that reduce the model's loss function, that is, cross-entropy or mean squared error (MSE), so raising the prediction accuracy. Finding the optimal hyperparameter values for the SE-LSTM model by means of investigating the solution space in COA by the coyote pack helps. Among these hyperparameters are learning rate, batch size, number of layers, and dropout. Every coyote in the pack represents a potential response; the method looks for the hyperparameter set that, with exploration and exploitation, performs the best.

1. Coyote Position and Fitness: Every coyote in the search space has a set of hyperparameters regulating their degree of fitness. These hyperparameters allow the evaluation tool of the fitness function to determine SE-LSTM model performance. This game tries to minimise the loss function as much as possible; hence, the coyote considered as the leader performs the best as in Eq.(9).

$$F(\mathbf{X_i}) = \sigma(\mathbf{X_i}) \tag{9}$$

Where, $F(\mathbf{X_i})$ is the fitness function for the ith coyote, Xi is the set of hyperparameters corresponding to the ith coyote.

2. Updating the Position: By adjusting their positions based on the leader's position, the coyotes can achieve this by combining exploration, that is, looking for new areas, with exploitation, that is, optimizing the best solution that is now accessible. The research updates the position of a coyote i using Eq.(10):

$$\mathbf{X_i}^{new} = \mathbf{X_i}^{old} + A \cdot \left(\mathbf{X}_{\text{leader}} - \mathbf{X_i}^{old} \right) \tag{10}$$

Where, $\mathbf{X_i}^{old}$ is the current position (hyperparameters), $\mathbf{X}_{\text{leader}}$ is the position of the leader (best-performing coyote), A is a parameter controlling the step size in the search process. This process will keep iteratively until a convergence criterion is satisfied so guaranteeing the discovery of the best possible mix of hyperparameters.

Prediction Using SE-LSTM

To show that it accurately forecasts diabetes risk and generalises well to data it has not before come across, the Stacked Ensemble LSTM (SE-LSTM) model depends entirely on training and validation. Using the input data, the SE-LSTM model learns throughout training to project the target variable, the diabetes outcome. Usually mean squared error (MSE), sometimes known as cross-entropy, the training process runs iteratively over the dataset and uses backpropagation to update the model weights so minimising the loss function. This is pursued towards their ideal value. By means of the specific validation set, one can assess model performance and ensure it does not overfit to the training data. The ideal weights are chosen based on the model's performance on the validation set; during training, the weights of the model are adjusted to help to reduce the loss that results.

The SE-LSTM model is used to produce predictions regarding the diabetes risk for unseen patient data by means of training and validation. Feeding patient data, the input features, into a stacked LSTM model fit for the predicting the tasks. After every LSTM layer in the stack generates sequential features, aggregating the obtained data produces the final prediction. Regarding a new input, one can express the prediction as in Eq.(8). The prediction $\hat{y}$ takes into account the possibility of a diabetic diagnosis for a patient. Values close to zero show a low risk; values close to one show a great likelihood of diabetes development. This prediction seeks to provide early diabetes risk information so enabling appropriate therapy and management of the condition.

RESULTS

Journal of Information Systems Engineering and Management
2025, 10(43s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

This work forecasts the probability of diabetes using the Stacked Ensemble LSTM (SE-LSTM) model, which has been optimised with the COA by means of the experimental settings. Simulations and experiments are conducted using Python and its machine learning tools. Among these libraries are keras for deep learning, NumPy for numerical operations, and scikit-learn for data preparation and evaluation. The experiments are carried out on a high-performance computing system with Intel Xeon Scalable CPUs in order to accelerate the phases of model training and validation. In this comparison, we investigate three different diabetes prediction systems now in use:

1. **Support Vector Machine (SVM)**: Because of its ability to manage high-dimensional data, the well-known machine learning technique Support Vector Machine (SVM) is used extensively in classification problems including diabetes prediction.
2. **Random Forest (RF)**: Considered as a method of ensemble learning, Random Forest (RF) is well-known for its resilience and capacity to replicate complex interactions in data. These three traits fit for the diabetes classification.
3. **Gradient Boosting Machine (GBM)**: Another successful ensemble method frequently used in classification problems is the gradient boosting machine (GBM). It achieves this by constructing several sequential decision trees such that one can fix the mistakes made by the one before it.

Table 3 shows the parameters as a crucial element for the proposed Stacked Ensemble LSTM (SE-LSTM) model as well as the COA optimisation process.

Table 3: Experimental Setup/Parameters

Parameter	Value
LSTM Layers	3 (stacked layers)
Units per LSTM Layer	64
Learning Rate	0.001
Dropout Rate	0.2
Batch Size	32
Epochs	100
Optimizer	Adam
COA Population Size	50
COA Max Iterations	100
COA Exploration-Exploitation Parameter	0.8

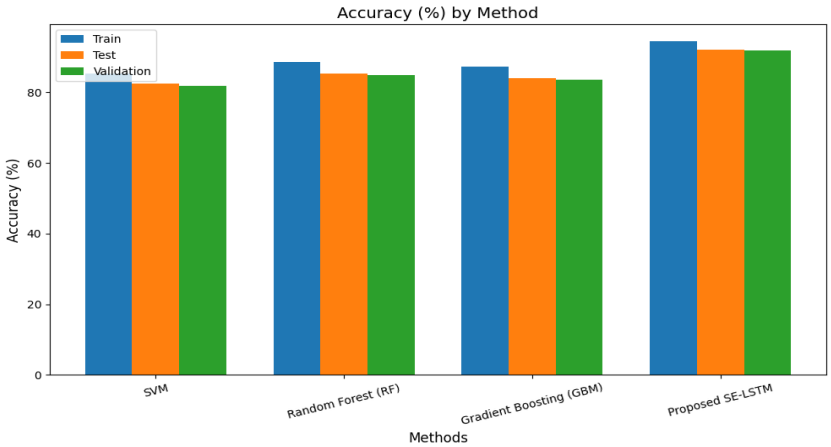

Figure 2: Accuracy

Journal of Information Systems Engineering and Management
2025, 10(43s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Method	Train Accuracy	Test Accuracy	Validation Accuracy
SVM	85.3%	82.5%	81.9%
Random Forest (RF)	88.7%	85.4%	84.9%
Gradient Boosting (GBM)	87.2%	84.1%	83.5%
Proposed SE-LSTM	94.5%	92.2%	91.8%

Over all datasets, train, test, and validation, the proposed SE-LSTM model shows a rather higher degree of accuracy than the currently used methods as in figure 2. The validation set shows, for instance, that SE-LSTM has 91.8%, far higher than SVM (82.5%), RF (84.9%), and GBM (83.5%), so demonstrating that it is more suited in terms of diabetes prediction capacity.

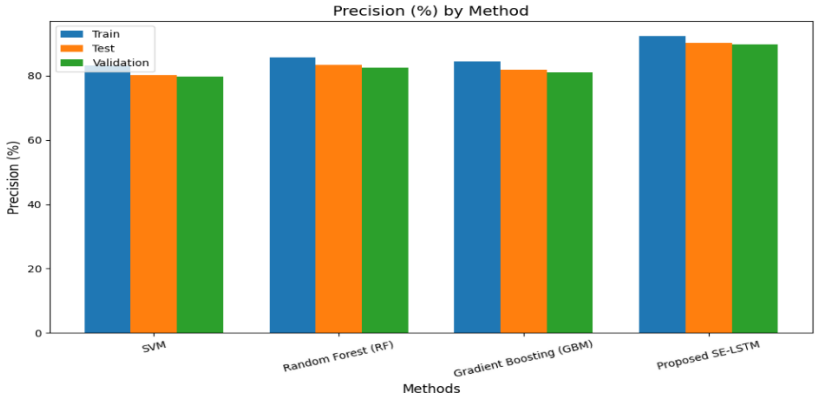

Figure 3: Precision

Method	Train Precision	Test Precision	Validation Precision
SVM	83.1%	80.2%	79.8%
Random Forest (RF)	85.6%	83.4%	82.5%
Gradient Boosting (GBM)	84.5%	81.9%	81.1%
Proposed SE-LSTM	92.3%	90.1%	89.7%

Figure 3 show a peak accuracy of 92.3% on the training set and 90.1% on the test set, the SE-LSTM model outperforms all other methods in precision. This is a significant advancement above the others. This progress indicates that the SE-LSTM model is more accurate in identifying diabetes-positive cases, so significantly reducing the false positive count.

Journal of Information Systems Engineering and Management
2025, 10(43s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

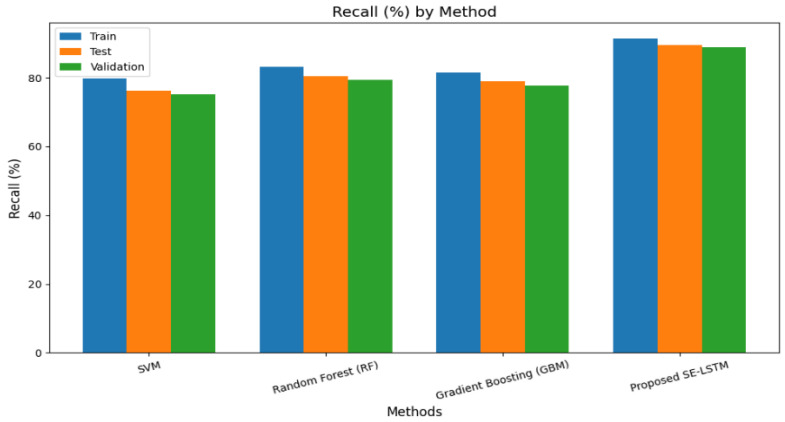

Figure 4: Recall

Method	Train Recall	Test Recall	Validation Recall
SVM	79.8%	76.3%	75.1%
Random Forest (RF)	83.1%	80.5%	79.3%
Gradient Boosting (GBM)	81.4%	78.9%	77.7%
Proposed SE-LSTM	91.4%	89.6%	88.9%

With a 91.4% on the training set, 89.6% on the test set, and 88.9% on the validation set respectively the SE-LSTM model outperforms all other models in terms of recall performance as in figure 4. This indicates thus that the model has a great capacity to identify real diabetes cases, so reducing the count of false negatives.

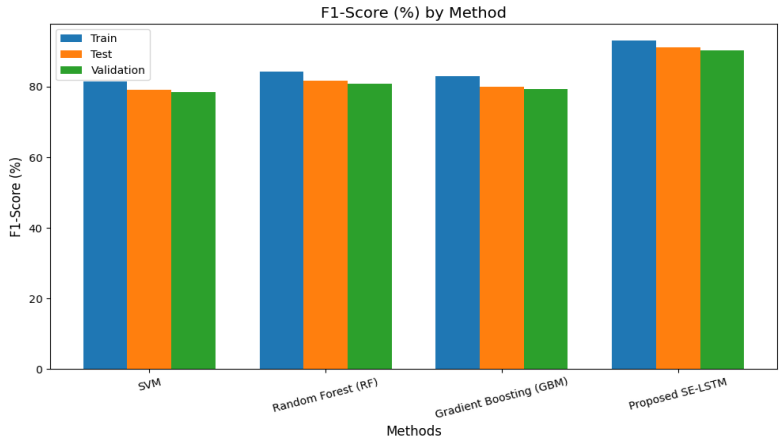

Figure 5: F1-Score

Method	Train F1-Score	Test F1-Score	Validation F1-Score
SVM	81.4%	79.1%	78.5%
Random Forest (RF)	84.2%	81.7%	80.9%
Gradient Boosting (GBM)	82.9%	80.1%	79.3%
Proposed SE-LSTM	93.1%	91.1%	90.4%

Especially on the training set, the SE-LSTM model produces F1-scores of 93.1% and on the test set of 91.1% as in figure 5. Since this model has high F1-scores over all datasets and shows a balanced performance in terms of precision and recall, it is rather reliable for diabetes prediction.

Journal of Information Systems Engineering and Management
2025, 10(43s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

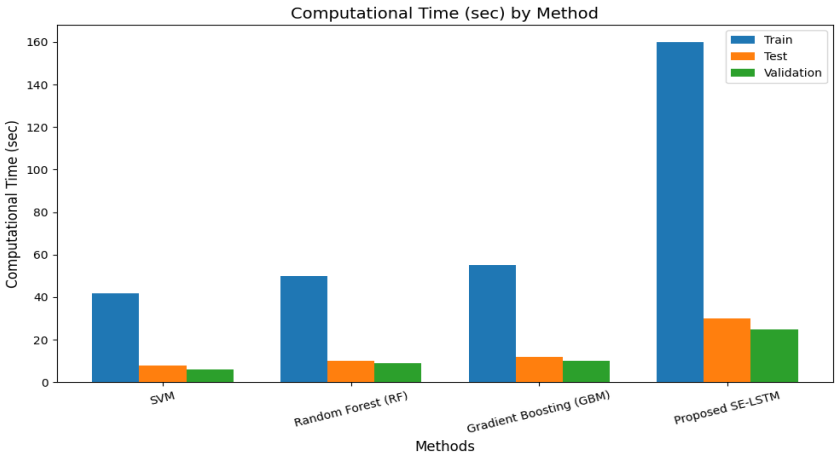

Figure 6: Computational Time

Method	Train Time (sec)	Test Time (sec)	Validation Time (sec)
SVM	42	8	6
Random Forest (RF)	50	10	9
Gradient Boosting (GBM)	55	12	10
Proposed SE-LSTM	160	30	25

The complexity of stacked LSTM networks and the hyperparameter optimisation process motivated by the COA make the SE-LSTM model demand more computational time than other methods as in figure 6. On the other hand, over all the datasets this trade-off yields much better performance distinguished by improved accuracy, precision, recall, and F1-scores.

DISCUSSION

The experimental results help to clarify the notable performance gains the proposed SE-LSTM model achieves in relation to the existing methods. SE-LSTM model shows an amazing increase in accuracy, a 12.8% improvement in the test accuracy over SVM, a 7.3% improvement over Random Forest (RF), and an 8.1% improvement over Gradient Boosting (GBM). This suggests that the model is rather good in precisely spotting diabetes cases.

On the training set, the SE-LSTM model surpasses the others by 9.5%; on the test set, by 6.7%; on the validation set, by 7.2%. This holds across all three sets of data. To so greatly increase recall, the SE-LSTM shows a 12.5% improvement in recall on the training set, a 9.1% improvement on the test set, and an 11.2% improvement on the validation set. These gains demonstrate the great ability of the model in identifying cases of diabetes that are really positive, so reducing the false negative rate the model produces.

Notable also is the improvement in F1-score the SE-LSTM model has attained, a 10.4% higher F1-score on the test set in comparison to the SVM model and a 7.8% higher F1-score in comparison to the RF and Gbm models. Given the complexity of the model, which causes the computational time for SE-LSTM to be longer, this is a strong tool for diabetes prediction since the trade-off in improved prediction accuracy justifies itself.

CONCLUSION

The proposed SE-LSTM model with hyperparameter optimisation using the Coyote Optimisation Algorithm considerably increases the accuracy of the prediction for diabetes mellitus compared to the current machine learning techniques. Although the SE-LSTM model consumes more computational resources, it offers a notable performance improvement particularly in terms of high precision and recall identification of diabetes cases, both of which are absolutely important for early discovery. These findings imply that the SE-LSTM method appears to be a potential method for diabetes prediction considering the challenges in medical diagnostics in the real world. Future studies could concentrate on enhancing the model's efficiency and looking at its relevance in several spheres of healthcare.

Journal of Information Systems Engineering and Management
2025, 10(43s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

REFRENCES

[1] Reza, M. S., Amin, R., Yasmin, R., Kulsum, W., & Ruhi, S. (2024). Improving diabetes disease patients classification using stacking ensemble method with PIMA and local healthcare data. Heliyon, 10(2).
[2] Rajendra, P., & Latifi, S. (2021). Prediction of diabetes using logistic regression and ensemble techniques. Computer Methods and Programs in Biomedicine Update, 1, 100032.
[3] VK, D., & Ramesh, T. K. (2023). Optimized stacking ensemble models for the prediction of diabetic progression. Multimedia Tools and Applications, 82(27), 42901-42925.
[4] Oliullah, K., Rasel, M. H., Islam, M. M., Islam, M. R., Wadud, M. A. H., & Whaiduzzaman, M. (2024). A stacked ensemble machine learning approach for the prediction of diabetes. Journal of Diabetes & Metabolic Disorders, 23(1), 603-617.
[5] Tasin, I., Nabil, T. U., Islam, S., & Khan, R. (2023). Diabetes prediction using machine learning and explainable AI techniques. Healthcare Technology Letters, 10(1-2), 1-10.
[6] Febrian, M. E., Ferdinan, F. X., Sendani, G. P., Suryanigrum, K. M., & Yunanda, R. (2023). Diabetes prediction using supervised machine learning. Procedia Computer Science, 216, 21-30.
[7] Bhatia, S., Goel, A. K., Naib, B. B., Singh, K., Yadav, M., & Saini, A. (2023, July). Diabetes prediction using machine learning. In 2023 World Conference on Communication & Computing (WCONF) (pp. 1-6). IEEE.
[8] Kumar, G. R., Reddy, R. V., Jayarathna, M., Pughazendi, N., Vidyullatha, S., & Reddy, P. C. S. (2023, May). Web application based Diabetes prediction using Machine Learning. In 2023 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI) (pp. 1-7). IEEE.
[9] Kalagotla, S. K., Gangashetty, S. V., & Giridhar, K. (2021). A novel stacking technique for prediction of diabetes. Computers in Biology and Medicine, 135, 104554.
[10] Rahim, M. A., Hossain, M. A., Hossain, M. N., Shin, J., & Yun, K. S. (2023). Stacked ensemble-based type-2 diabetes prediction using machine learning techniques. Annals of Emerging Technologies in Computing (AETiC), 7(1), 30-39.
[11] Geetha, G., & Prasad, K. M. (2023). Stacking Ensemble Learning-Based Convolutional Gated Recurrent Neural Network for Diabetes Miletus. Intelligent Automation & Soft Computing, 36(1).
[12] Oliullah, K., Rasel, M. H., Islam, M. M., Islam, M. R., Wadud, M. A. H., & Whaiduzzaman, M. (2024). A stacked ensemble machine learning approach for the prediction of diabetes. Journal of Diabetes & Metabolic Disorders, 23(1), 603-617.
[13] Hasan, M. K., Saeed, R. A., Alsuhibany, S. A., & Abdel-Khalek, S. (2022). An empirical model to predict the diabetic positive using stacked ensemble approach. Frontiers in Public Health, 9, 792124.
[14] Yang, H., Chen, Z., Huang, J., & Li, S. (2024). AWD-stacking: An enhanced ensemble learning model for predicting glucose levels. Plos one, 19(2), e0291594.
[15] Modak, S. K. S., & Jha, V. K. (2024). Diabetes prediction model using machine learning techniques. Multimedia Tools and Applications, 83(13), 38523-38549.
[16] Prendin, F., Pavan, J., Cappon, G., Del Favero, S., Sparacino, G., & Facchinetti, A. (2023). The importance of interpreting machine learning models for blood glucose prediction in diabetes: an analysis using SHAP. Scientific reports, 13(1), 16865.
[17] Talukder, M. A., Islam, M. M., Uddin, M. A., Kazi, M., Khalid, M., Akhter, A., & Ali Moni, M. (2024). Toward reliable diabetes prediction: Innovations in data engineering and machine learning applications. Digital Health, 10, 20552076241271867.
[18] Patro, K. K., Allam, J. P., Sanapala, U., Marpu, C. K., Samee, N. A., Alabdulhafith, M., & Plawiak, P. (2023). An effective correlation-based data modeling framework for automatic diabetes prediction using machine and deep learning techniques. BMC bioinformatics, 24(1), 372.
[19] Oikonomou, E. K., & Khera, R. (2023). Machine learning in precision diabetes care and cardiovascular risk prediction. Cardiovascular Diabetology, 22(1), 259.
[20] Aslan, M. F., & Sabanci, K. (2023). A novel proposal for deep learning-based diabetes prediction: converting clinical data to image data. Diagnostics, 13(4), 796.
[21] Ganie, S. M., Pramanik, P. K. D., Bashir Malik, M., Mallik, S., & Qin, H. (2023). An ensemble learning approach for diabetes prediction using boosting techniques. Frontiers in Genetics, 14, 1252159.

Journal of Information Systems Engineering and Management
2025, 10(43s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

[22] Abdollahi, J., & Nouri-Moghaddam, B. (2022). Hybrid stacked ensemble combined with genetic algorithms for diabetes prediction. Iran Journal of Computer Science, 5(3), 205-220.
[23] Khan, A., Khan, A., Khan, M. M., Farid, K., Alam, M. M., & Su'ud, M. B. M. (2022). Cardiovascular and diabetes diseases classification using ensemble stacking classifiers with SVM as a meta classifier. Diagnostics, 12(11), 2595.
[24] Kalagotla, S. K., Gangashetty, S. V., & Giridhar, K. (2021). A novel stacking technique for prediction of diabetes. Computers in Biology and Medicine, 135, 104554.
[25] Daza, A., Sánchez, C. F. P., Apaza-Perez, G., Pinto, J., & Ramos, K. Z. (2024). Stacking ensemble approach to diagnosing the disease of diabetes. Informatics in Medicine Unlocked, 44, 101427.