Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

A Hybrid LSTM-Autoencoder Framework for Accurate Prediction of Diabetic Kidney Disease

Konne Madhavi[1,*], Harwant Singh Arri[2]
[1,2]*Department of Computer Science and Engineering, Lovely Professional University, Phagwara, Punjab, India*
E-mail: konnemadhavi@gmail.com, hsarri@gmail.com

ARTICLE INFO

Received: 26 Dec 2024

Revised: 14 Feb 2025

Accepted: 22 Feb 2025

ABSTRACT

Diabetic Kidney Disease (DKD), a complication of diabetes, leads to the gradual decline of renal function. Screening individuals for DKD is essential, as a timely intervention has been shown to improve patient outcomes. Although early detection of DKD can prevent its progression, systematic screening is not universally feasible, which can result in missed or delayed diagnoses. Deep learning (DL) models have shown promise in the medical field, providing promising results. Many researchers have proposed DL models for DKD classification, but achieving reliable accuracy has still been challenging due to the presence of noise and unwanted features in medical data. To address this issue, the paper proposes a hybrid DL model by integrating Long Short-Term Memory (LSTM) into an Autoencoder (AE) architecture, called LSTM-AE, for effective DKD prediction. The combination of these two models effectively identifies important features, resulting in accurate classification. The data used to analyze the DL model's performance was collected from the UCI repository. The data was affected by various issues, and several preprocessing steps were performed to clean the data. This preprocessing also contributed to achieving effective outcomes. The proposed model was compared with three popular DL models: Convolutional Neural Network (CNN), LSTM, and AE. The LSTM-AE achieved the highest accuracy of 99% in DKD prediction, while the other models produced accuracies ranging from 94% to 97%. The proposed model was also compared with existing models from recent studies, and in all experimental outcomes, the LSTM-AE outperformed the others. The results demonstrate that the proposed model is reliable for DKD prediction and can be deployed in real-time practice.

Keywords: Diabetic Kidney Disease, Hybrid Deep Learning, Autoencoder, Long Short-Term Memory, UCI Repository

INTRODUCTION

DKD is one of the most severe problem of diabetes [1]. Traditional thought held that DKD was distinct from macrovascular disease (peripheral vascular disease, coronary heart disease, and cerebrovascular disease) and was more closely related to microvascular disorders (neuropathy and retinopathy) [2]. Nonetheless, each disease can be viewed as a specific tissue's expression of a common pathogenetic process; in this case, DKD is the kidney's expression of a glucose-driven process that occurs in other sensitive areas of the body. Microalbuminuria referred to as occult or incipient nephropathy, is a degenerative condition characterized by a gradual increase in urinary albumin excretion (30-300 mg/day), originally described by Mogensen in 1980 regarding DKD [3]. When albuminuria became identifiable with conventional dipstick urinalysis (>300 mg/day), macroalbuminuria was introduced to denote the escalating quantity of albumin excreted in the urine.

DKD affects 20-40% of Chinese adults and is the primary cause of last-stage renal disease in the country [4] . This disorder has a terrible effect on diabetes sufferers' quality of life. Patients, their families, and society as a whole bear a disproportionate amount of the financial costs involved with DKD therapy. Microalbuminuria levels are commonly used in the early clinical stages of DKD. Unfortunately, by the time DKD is diagnosed, irreversible damage to the kidney has already occurred[5]. Even with aggressive treatment, the condition will worsen for more than one-third of patients. The implementation of preventative interventions for all diabetic patients will require substantial human resources and medical expenses. Considering the low incidence of DKD in certain diabetic patients, it is essential to

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

identify those at high risk for future preventative interventions, thus demanding the implementation of advanced screening technology [6] [7].

Early detection of DKD can save a patient's life. To accurately diagnose kidney problems, medical practitioners use a variety of standard approaches, including physical examinations and laboratory tests [8]. Building trustworthy and generalizable diagnostic models that may assist medical practitioners in making swift and informed judgments is critical, especially given new potential data sources that can help medical diagnosis.

In the realm of medical diagnosis, machine learning (ML) has lately aided in the development of efficient models capable of reaching quick and correct results. DL is a subset of ML that employs a sequence of operations conducted during training to uncover hidden correlations in a dataset [9]. Medical applications are heavily influenced by DL, a multilayer DL model that can potentially handle nonlinear data. Medical data heterogeneity presents a variety of issues for DL in terms of generalizability and robustness, potentially leading to erroneous rules and unreproducible diagnostic models unless DL becomes more widely used [10]. As a result, even if ideal weights are reached during DL training, the model may still have significant variance. Using a range of DL models could help overcome this challenge. The term for this strategy is hybrid learning. By combining the benefits of more than one DL model, we may overcome the limitations of single models and obtain higher generalizability and flexibility [11, 12]. Multiple studies [13, 14] have demonstrated that hybrid learning yields accurate and efficient models. The most crucial aspect of developing a decent model is selecting the appropriate set of features. The DL field has done extensive research on feature selection, with promising results in biomedical applications [15]. Based on the information supplied, the primary purpose of this work is to develop a hybrid DL model that can improve DKD prediction performance by using the best features from the dataset.

The contribution of the research is as follows:

- Develop a hybrid DL model, LSTM-Autoencoder, for effective feature extraction from the data and to improve DKD prediction accuracy.
- A detailed exploratory data analysis and visualization are conducted to clean the data, which helps improve the proposed model's performance.
- The proposed hybrid DL model is evaluated using positive and negative metrics, in comparison with other single DL models such as CNN, LSTM, and AE.
- A comparison with existing studies also demonstrates the promising results of the proposed model.

The research paper is organized in the following manner: Section 1 discusses DKD, traditional identification methods, and the need for integrating AI in DKD prediction. Section 2 discusses recent related works in kidney disease detection. Section 3 provides the architecture and working of the proposed hybrid LSTM-AE model. Section 4 presents the experimental setup, results, and their related discussion, and also covers data acquisition and preprocessing. Finally, the research concludes with a discussion of future work in Section 5.

LITERATURE SURVEY

The literature discusses and applies a range of approaches to the DKD categorization problem. Given the current literature, the proposed study contributes to the field's efforts to improve the discoveries that are now possible in DKD prediction. The paper [16] introduces a novel AI method for estimating chronic kidney disease (CKD), incorporating numerous pre-process steps, feature selection, and hyperparameter optimization. To fill in missing values in the dataset, a new sequential data scaling method and iterative imputation are employed. The Boruta approach is utilized for feature selection, while ML methods are employed to build the model. During validation on the UCI dataset, the proposed model performed quite well. The strategy, which uses unique preprocessing procedures, enhances the early detection of CKD. As proven in this work, ML methods can improve clinical support systems as well as the role of uncertainty in chronic disease prognosis. A bagging classifier and voting algorithms are employed to combine the DL methodology known as artificial neural network (ANN) with the CKD dataset available from the UCI. The dataset contains four hundred cases and twenty-five features. This work is described in article [17]. Three EDL models are constructed as a foundation for the study: Model-1, which integrates an ANN with a

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

voting method; Model-2, which incorporates a bagging method; and Model-3, which integrates both methods. Considering the experimental metrics for performance, it is evident that the suggested model surpasses all previously examined methodologies.

The research study [18] provides a new model for early detection and prediction of CKD that employs a hybrid DL network. This paper presents a DL strategy for early CKD detection based on a Deep Separable Convolution Neural Network (DSCNN). To detect kidney abnormalities, the Capsule Network pulls additional processing features from certain attributes. The Aquila Optimization Algorithm is used to swiftly select the most relevant attributes for categorization. With the correct features, categorization becomes more successful and requires less computational power. The Sooty Tern Optimization Algorithm is employed to enhance the DSCNN method for diagnosing CKD in patients. The dataset is subsequently evaluated using the CKD dataset from the UCI ML repository. The experimental results indicate that the proposed strategy surpasses the current state-of-the-art method for classifying CKD.

In the study [19], it was demonstrated using huge data that AI could forecast the progression of DKD. The study used three distinct strategies to improve the capacity to predict disease-specific outcomes. A new diabetic complications prediction model was developed before patients exhibiting clinical signs such as microalbuminuria. Second, instances that were not clinically defined as type 2 diabetes in their EMR text were included, and massive amounts of data from electronic medical records were used for ML, with no intention of performing clinical trials. Third, to estimate DKD progression for the six months after the reference periods, AI used time-series data obtained six months before those periods. An AI-powered predictive model for DKD progression detection may result in more precise and efficient intervention, reducing the requirement for hemodialysis. The goal of this study [20] is to offer a new ensemble DL method for detecting CKD; multiple feature selection approaches were used to choose the optimal features. Furthermore, from a medical standpoint, the study looks into how the best features selected affect CKD. The proposed ensemble model employs a support vector machine as the meta-learner model and includes pre-trained DL models. Extensive testing was conducted using UCI repository data. The findings indicate that the proposed model is more effective than existing models in predicting CKD. The suggested model with the given features produced the best results.

PROPOSED LSTM-AE MODEL

In this section, we first provide a general architecture of AE and LSTM, and the proposed hybrid LSTM-AE architecture is explained with a block diagram.

Autoencoder

AE, a significant type of neural network, exhibits characteristics that render them advantageous for various applications [21]. **Figure 1** illustrates the architecture of the AE, comprising an encoder, a bottleneck layer, and a decoder [22].

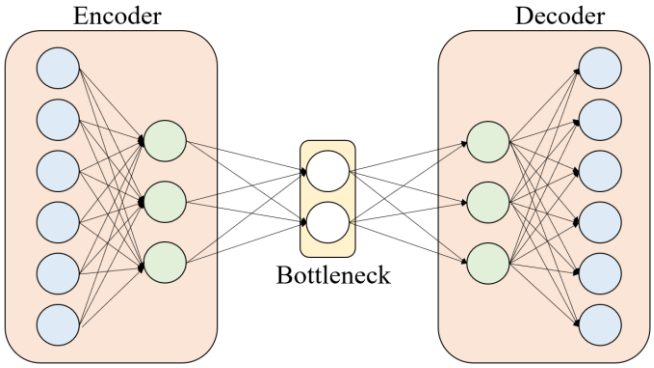

Figure 1. Autoencoder architecture, comprising an encoder, bottleneck layer, and decoder.

The AE architecture compels data to traverse the bottleneck layer, which has fewer nodes than the input layer, to achieve the objective of reconstructing the input [23]. This architectural characteristic compels the encoder to

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

identify the essential hidden attributes required for supplying the decoder with reconstruction data from the input [24]. Equation (1) is employed to define L as the AE reconstruction.

$$L = D\big(E(X)\big) \quad (1)$$

$$L_a = D\big(E(X_{dkd})\big) \quad (2)$$

$$L_b = D\big(E(X_{non-dkd})\big) \quad (3)$$

When the initial input is represented by X, the encoder by E, and the decoder by D, Equations (2) and (3) are obtained from subscripted versions of X and L that are specific to the investigation. X_{dkd} and L_{dkd} are used exclusively for DKD data, while $X_{non-dkd}$ and $L_{non-dkd}$ are used only for non-DKD data. AE training seeks to reduce the disparity between the actual input and the reconstructed output. This is achieved by computing the equation $l(X - L)$, where l represents a loss function. According to the notation, when inferring from new samples, $L_{non-dkd}$ contains more data than L_{dkd}. AE feature residuals are the differences between an AE's initial input and its reconstruction, as studied independently for each feature. As illustrated in Equation (4), the AE feature residuals are defined as S.

$$S = X - L \quad (4)$$

$$S_{dkd} = X_{dkd} - L_{dkd} \quad (5)$$

$$S_{non-dkd} = X_{non-dkd} - L_{non-dkd} \quad (6)$$

In this scenario, X represents the original input, and L is the AE reconstruction stated in Equation (1). The same subscript notation for S is used as for X and L. S_{dkd} focuses solely on DKD data, whereas $S_{non-dkd}$ focuses solely on non-DKD data, yielding Equations (5) and (6). When discussing S in this study, it is critical to note that a summary statistic for each group is not employed. The number of dimensions in S remains the same as in X. This ensures that each feature in a sample has a residue. It is expected that $L_{non-dkd}$ will be larger than L_{dkd} after pretraining with only $X_{non-dkd}$. As a result, $S_{non-dkd}$ is expected to have fewer data points than S_{dkd}, as they are directly related to reconstruction performance.

An alternative perspective on S is that it retains the parts of X that are difficult to reconstruct, whereas L holds the parts that are easier to reconstruct. Equations (7)–(9) reinforce this way of thinking by demonstrating that L and S are distinct portions of X.

$$X = L + S \quad (7)$$

$$X_{dkd} = L_{dkd} + S_{dkd} \quad (8)$$

$$X_{non-dkd} = L_{non-dkd} + S_{non-dkd} \quad (9)$$

Using this strategy, the following theoretical conditions can be derived for obtaining the best features to get the best classification: flawless reconstruction of non-DKD data; and an inability to reconstruct DKD data. All non-DKD data exhibit zero AE feature residuals, while for DKD data, the input matches the AE feature residuals.

LSTM

The LSTM network is a common variation of the RNN that can learn long-term dependencies and address the vanishing gradient issue. It was first introduced in 1997 [25] but has only recently gained prominence due to its widespread application in commercial contexts. Training with backpropagation over time often results in the vanishing or expanding gradient issue in RNN architectures [26]. As the calculation progresses through the network, the computed derivatives during training can expand or shrink exponentially. In extreme cases, the model either stops training entirely when the gradient diminishes or trains erratically, failing to minimize loss. The RNN weight values remain either unchanged or are significantly altered. **Figure 2** depicts a visualization of LSTM model.

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

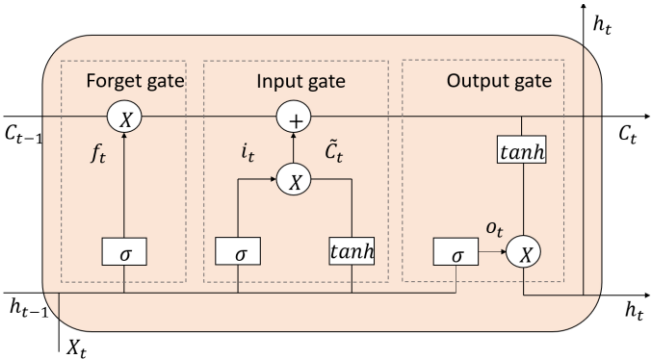

Figure 2: Architecture of the LSTM, employs forget, input, and output gates to regulate the flow of data and address the vanishing gradient issue. These gates enable the model to learn long-term dependencies by controlling what data is forgotten, retained, or passed on to subsequent time steps.

The LSTM architecture resolves this issue by incorporating "gates" into the design, enabling the cell state (CS) to forget data, replace values, and determine which could be output and transferred to the other cell [27]. The input x_t and the past hidden CS h_{t-1} are fed into the forget gate's sigmoid layer, which produces the vector f_t. This vector specifies which irrelevant data from C_{t-1} should be removed from the current CS.

$$f_t = \sigma\left(W_f * [h_{t-1}, x_t] + b_f\right) \quad (10)$$

The input gate regulates the information retained in the CS by employing a sigmoid on the x_t and the h_{t-1} to ascertain which values to modify, i_t. A $tanh$ produces candidate values, $\tilde{C}_t$, for updating the CS. The product of these vectors indicates the values to be integrated into the new CS.

$$i_t = \sigma(W_i * [h_{t-1}, x_t] + b_i) \quad (11)$$

$$\tilde{C}_t = tanh\,(W_C * [h_{t-1}, x_t] + b_C) \quad (12)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (13)$$

The output gate decides which portion of the preceding hidden CS h_{t-1} to transmit to the subsequent cell by employing a sigmoid function on the x_t and the h_{t-1}. Subsequently, it computes the product of the hyperbolic tangent of C_t and o_t.

$$o_t = \sigma(W_o * [h_{t-1}, x_t] + b_o) \quad (14)$$

$$h_t = o_t * tanh(C_t) \quad (15)$$

Hybrid LSTM-AE

The architecture of the LSTM and AE models was discussed earlier. To improve the accuracy of DKD detection, we propose the LSTM-AE model. **Figure 3** presents a detailed overview of this architecture. In the LSTM-AE model, the traditional encoder and decoder of the AE model are replaced by LSTM networks to process the data. The data is input into the network, with the encoder consisting of three layers: an input layer, a normalization layer (BatchNorm) [28], and an LSTM hidden layer, which produces a two-dimensional real vector sequence. The input layer accepts the data, and the BatchNorm layer normalizes it. The LSTM layers in the encoder then condense the data into a compact representation, which is output as a two-dimensional vector sequence. The decoder receives the encoder's output and is symmetrically designed to mirror the encoder's structure. It processes the data through three concealed layers: input, normalization, and LSTM. The LSTM blocks in both the encoder and decoder help condense and reconstruct the features by operating in the latent space. Temporal information is processed horizontally, with the LSTM layers being reused multiple times to refine the features. Spatial information is processed vertically, where the output of each hidden state serves as the input for the next layer. The depth of the network can be increased to enhance its performance, refining the fit between the input and output. The LSTM cells are particularly beneficial for preserving long-term dependencies in the data, eliminating the need for exponential decay patterns. The final output

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

of the decoder is fed into a fully connected layer, which extracts the key features from the LSTM-AE model and performs classification through two output neurons.

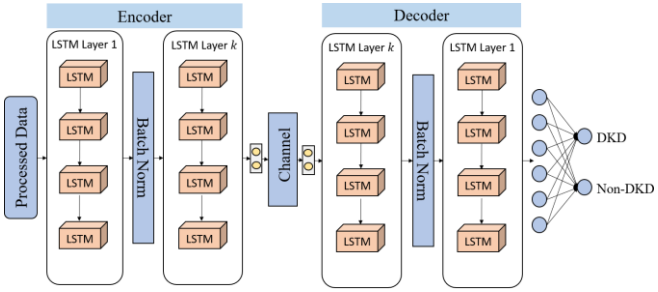

Figure 3: Proposed LSTM-AE Architecture replaces the encoder and decoder of an AE with LSTM layers. The LSTM-AE compresses data into a 2D latent space via the encoder, reconstructs it symmetrically through the decoder, and performs classification using a fully connected layer.

EXPERIMENTAL RESULT AND ANALYSIS

This section details the experimental setup, including the programming language, software used, and hardware configuration, as well as the data acquisition and processing. The performance of the proposed model is compared with traditional models for DKD prediction. Additionally, a comparison with recent works from journals is also provided.

Experimental setup and data description

For DKD prediction, Jupyter Notebook [29] and Python programming language were chosen. The system configuration includes an Intel Core i5-9500T CPU at 2.20 GHz, 8 GB RAM, and a 64-bit operating system.

Data Acquisition: In this investigation, the dataset used to train the models was obtained from the UCI ML repository [30] . Classification attributes, such as DKD and non-DKD, indicate the patient's DKD status. This dataset contains 400 entries with 25 features, including 14 numerical and 11 categorical features. **Figure 4** provides additional details of the data, including a comprehensive summary of the attributes analyzed in the study.

id	157	109	17	347	24
age	62	54	47	43	42
bp	70	70	80	60	100
sg	1.025			1.025	1.015
al	3			0	4
su	0			0	0
rbc	normal			normal	normal
pc	abnormal			normal	abnormal
pcc	notpresent	notpresent	notpresent	notpresent	notpresent
ba	notpresent	notpresent	notpresent	notpresent	present
bgr	122	233	114	108	
bu	42	50.1	87	25	50
sc	1.7	1.9	5.2	1	1.4
sod	136		139	144	129
pot	4.7		3.7	5	4
hemo	12.6	11.7	12.1	17.8	11.1
pcv	39			43	39
wc	7900			7200	8300
rc	3.9			5.5	4.6
htn	yes	no	yes	no	yes
dm	yes	yes	no	no	no
cad	no	no	no	no	no
appet	good	good	poor	good	poor
pe	no	no	no	no	no
ane	no	no	no	no	no
classification	dkd	dkd	dkd	notdkd	dkd

Figure 4: DKD dataset Sample. The dataset contains 400 entries with 25 features, including 14 numerical and 11 categorical attributes, which describe various factors related to the DKD and non-DKD.

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Each feature in the dataset is analyzed for better understanding. The distribution of numerical and categorical features is examined using plots, as shown in **Figures 5 and 6**. This analysis helps to gain insights into each feature, such as its maximum and minimum values.

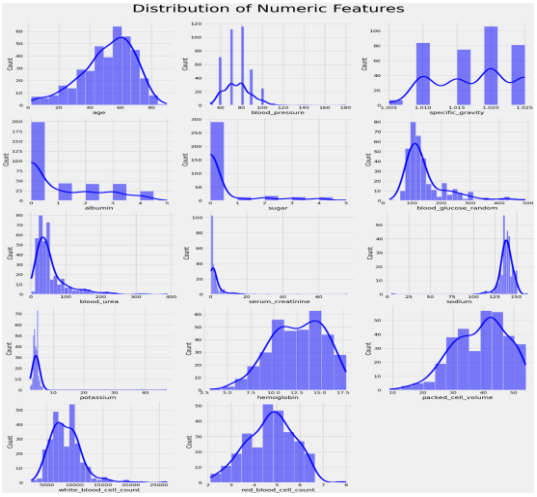

Figure 5: Distribution of numerical features in the DKD dataset shows the spread of values for each numerical attribute, providing information on the range, tendency, and outliers.

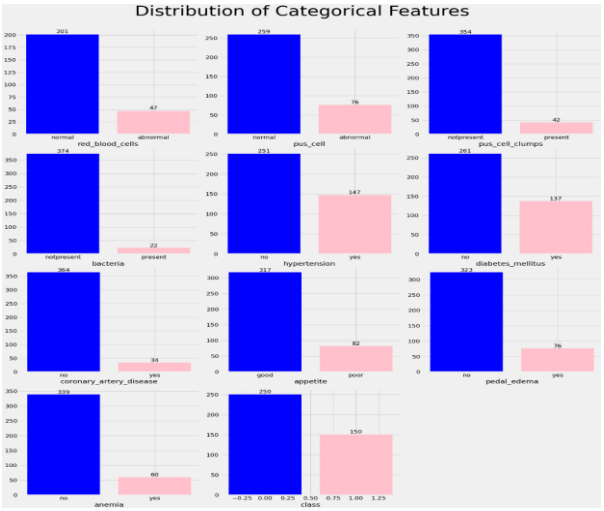

Figure 6: Distribution of categorical features in the DKD dataset, illustrating the frequency distribution of categorical attributes.

Data Processing:

Medical datasets are prone to numerous issues that might adversely affect the efficiency of the DL models [31]. Consequently, tackling these difficulties is crucial for improving data quality. The pre-processing phase is essential for enhancing data quality by addressing fundamental concerns such as handling missing data, category-to-numeric conversion, normalization, and outlier removal.

Handling missing data:

This dataset contained a substantial amount of missing values. The missing data in numerical and categorical features is visually depicted in **Figures 7 and 8.** Addressing missing data necessitates the selection of suitable statistical methods depending on the degree of missing values and the importance of the excluded variable. Conventional methods are effective when the proportion of missing values is low. To resolve this issue, iterative

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

imputation [32] was employed. This statistical method predicts missing values sequentially using observed data while considering the interrelationships among variables.

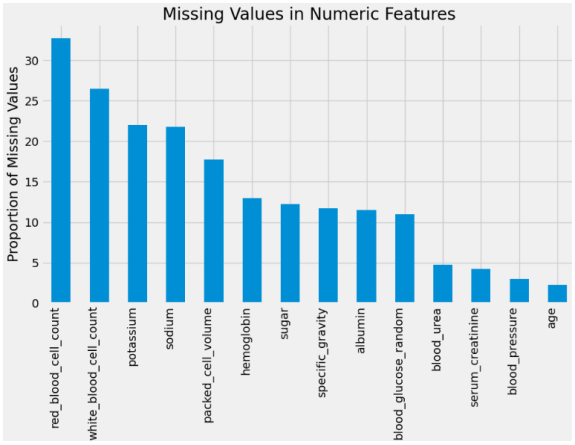

Figure 7: Analysis of missing values in the numerical features of the DKD dataset. This figure shows that the "red_blood_cell_count" feature has the most missing values, while the "age" feature has the least.

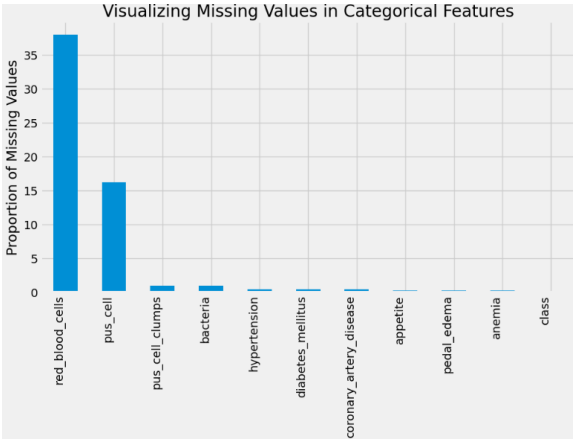

Figure 8: Analysis of missing values in the categorical features of the DKD dataset. This figure shows that the "red_blood_cell" feature has the most missing values, while the "class" feature has the least.

Handling categorical data: The dataset included both numerical and categorical features, which were managed using the label encoder module [33]. This module enhanced the performance of the DL model by transforming categorical features into numerical representations.

Handling imbalanced data

An oversampling approach was employed to achieve dataset balancing. A SMOTE-based oversampling strategy was utilized for this purpose, avoiding the random generation of data points [34]. Specifically, k closest neighbors from the minority class were selected at random. A neighbor's vector was identified as the one connecting the selected data point. This vector was scaled by a random value between 0 and 1, and the resulting synthetic data point was generated by combining the scaled vector with the original data point. After dataset balancing, the next step involved eliminating any noise, including outliers.

Handling outliers

Outliers [35], which are feature values significantly deviating from the average range, present a critical challenge in building a reliable and meaningful model. To address this, a statistical analysis of all data was performed to detect outliers and evaluate their medical relevance. Outliers in the dataset were replaced using the feature mean.

Normalization:

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Min-max scaling [36] was employed to normalize the features within a specified range, often between 0 and 1, by subtracting the minimum value (x_{min}) and dividing by the range ($x_{max} - x_{min}$). Min-max scaling operation is given in Equation

$$Min - Max\ Scaling(x) = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (16)$$

After completing all the above-mentioned steps, the data is processed and ready to be fed into the DL model.

EXPERIMENTAL RESULT

The processed data is fed into the proposed LSTM-AE model for DKD prediction. To evaluate the performance of the LSTM-AE model, three other popular DL models—CNN, LSTM, and AE—are also utilized. For all models, the same configuration is maintained, including the number of neurons and activation functions in the output layer, optimizer, loss function, epochs, batch sizes, and training and testing samples. The outcomes of the LSTM-AE model and other DL models are presented in Table 1. For evaluation, positive metrics such as accuracy, precision, F1-score, true positive rate (TPR), and true negative rate (TNR) are utilized, along with negative metrics such as false positive rate (FPR) and false negative rate (FNR) [37, 38]. The formulas used to calculate these metrics are provided in Equations (17–23).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (17)$$

$$Precision = \frac{TP}{TP+FP} \quad (18)$$

$$F1 - Score = \frac{TP}{TP+\frac{1}{2}(FN+FP)} \quad (19)$$

$$TNR = \frac{TN}{TN+FP} \quad (20)$$

$$TPR = \frac{TP}{TP+FN} \quad (21)$$

$$FNR = \frac{FN}{TP+FN} \quad (22)$$

$$FPR = \frac{FP}{FP+TN} \quad (23)$$

In the above equations, TP (True Positive) represents the number of correctly detected DKD samples. TN (True Negative) represents the total number of correctly detected non-DKD samples. FP (False Positive) and FN (False Negative) represent the number of wrongly identified DKD and non-DKD samples, respectively.

The performance metrics in Table 1 indicate that the LSTM-AE model achieves the highest accuracy of 99%, outperforming CNN (94%), LSTM (96%), and AE (97%). This demonstrates the model's ability to correctly predict a greater number of samples. For precision and F1-score, the LSTM-AE model again outperforms the other models, achieving scores of 98% and 98.99%, respectively. The TPR and TNR values of the LSTM-AE model are 100% and 98.04%, significantly higher compared to the other three models. Additionally, the proposed model achieves an FNR of 0%, meaning it never misses any DKD samples, and an FPR of 1.96%, indicating minimal false positives. The table highlights the excellence of the proposed model in identifying important features from the dataset and delivering highly accurate classifications. The comparison of positive and negative metrics of the proposed model with conventional DL models is illustrated in **Figures 9 and 10.**

TABLE I. PERFORMANCE METRICS OF DL MODEL FOR DKD PREDICTION

Model	Accuracy	Precision	F1-Score	TNR	TPR	FNR	FPR
CNN	94.00	94.12	94.12	93.88	94.12	5.88	6.12
LSTM	96.00	96.08	96.08	95.92	96.08	3.92	4.08
AE	97.00	97.96	96.97	98.00	96.00	4.00	2.00
Proposed LSTM-AE	99.00	98.00	98.99	98.04	100.00	0.00	1.96

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

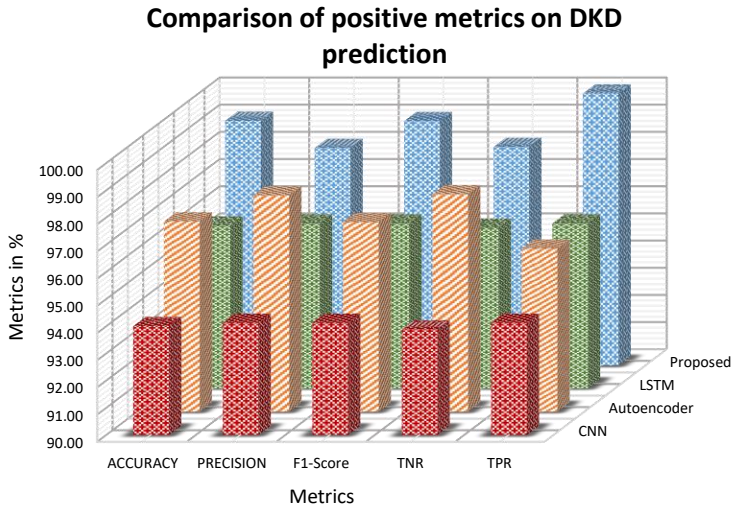

Figure 9: Evaluation of positive metrics for DKD prediction using the LSTM-AE, CNN, LSTM, and AE models.

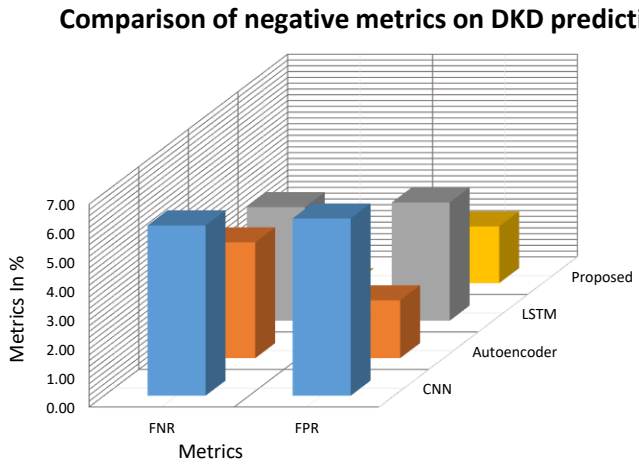

Figure 10: Evaluation of negative metrics for DKD prediction using the LSTM-AE, CNN, LSTM, and AE models.

A. State-of-art Comparison

The proposed LSTM-AE model is compared with recent works in existing research from 2022 to 2024. Among the reviewed studies, the highest accuracy achieved was by a Random Forest model, which reached 98.45%. However, our proposed model achieves an accuracy of 99% in DKD prediction, surpassing the best-performing model by 0.55%. Table 2 provides a comparison of the proposed method with existing work, highlighting the promising results of the proposed model in DKD prediction. While most studies focused solely on positive metrics for evaluation, this study considers both positive and negative metrics to comprehensively analyze the performance of the proposed model.

TABLE II. COMPARISON WITH EXISTING METHODS

Ref	Model	Accuracy	Precision	F1-Score	TNR	TPR
Our	Proposed LSTM-AE	99.00	98.00	98.99	98.04	100.00
[39]	Random Forest	86	83.6	-	-	86.2

[40]	Gradient Boosting	97	-	-	98	99
[41]	Bagging Tree with optimization	96.29	98.78	97	98	95.29
[42]	Random Forest	98.45	-	-	-	-

CONCLUSION

The research successfully designs a hybrid DL model, LSTM-AE, for DKD prediction, achieving a maximum accuracy of 99%. Early prediction of DKD is crucial for saving human lives. However, current research struggles to provide accurate DKD predictions, even when using AI techniques. The reasons for this include the highly noisy, imbalanced, and complex nature of medical data. This research offers a solution to these problems. In this study, the strengths of two models, AE and LSTM, are combined, resulting in improved performance. The proposed model is evaluated against CNN, LSTM, and AE models, demonstrating its effectiveness through various metrics. The proposed model achieves the highest scores for Accuracy, Precision, F1-Score, TNR, and TPR, with values of 99.00%, 98.00%, 98.99%, 98.04%, and 100.00%, respectively. It also achieves the lowest scores for FNR and FPR at 0.00% and 1.96%, respectively.

The experimental results are promising. The research focuses on using a single dataset, but for real-world deployment, testing with diverse datasets is crucial. Further external datasets should also be used to validate the model. The model will eventually be made accessible to healthcare professionals and patients through an application, allowing anyone to access it via mobile or laptop.

ACKNOWLEDGEMENTS

The authors are thankful to the staff and department of Computer Science and Engineering, Lovely Professional University for their support

REFERENCES

[1] Hoogeveen EK. The Epidemiology of Diabetic Kidney Disease. 2022;2(3):433-42.
[2] Li Y, Liu Y, Liu S, Gao M, Wang W, Chen K, et al. Diabetic vascular diseases: molecular mechanisms and therapeutic strategies. Signal Transduction and Targeted Therapy. 2023;8(1):152.
[3] Ceccarelli Ceccarelli D, Paleari R, Solerte B, Mosca A. Re-thinking diabetic nephropathy: Microalbuminuria is just a piece of the diagnostic puzzle. Clinica chimica acta; international journal of clinical chemistry. 2022;524:146-53.
[4] Trillini M, Perico N, Remuzzi G. Chapter 1 - Epidemiology of End-Stage Renal Failure: The Burden of Kidney Diseases to Global Health. In: Orlando G, Remuzzi G, Williams DF, editors. Kidney Transplantation, Bioengineering and Regeneration: Academic Press; 2017. p. 5-11.
[5] Hoogeveen EK. The Epidemiology of Diabetic Kidney Disease. Kidney and Dialysis [Internet]. 2022; 2(3):[433-42 pp.].
[6] George C, Echouffo-Tcheugui JB, Jaar BG, Okpechi IG, Kengne AP. The need for screening, early diagnosis, and prediction of chronic kidney disease in people with diabetes in low- and middle-income countries-a review of the current literature. BMC medicine. 2022;20(1):247.
[7] Wang N, Zhang C. Recent Advances in the Management of Diabetic Kidney Disease: Slowing Progression. International Journal of Molecular Sciences [Internet]. 2024; 25(6).
[8] Chen TK, Knicely DH, Grams ME. Chronic Kidney Disease Diagnosis and Management: A Review. Jama. 2019;322(13):1294-304.
[9] Sarker I. Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. SN Computer Science. 2021;2.

[10] Lee CH, Yoon HJ. Medical big data: promise and challenges. Kidney research and clinical practice. 2017;36(1):3-11.
[11] Gavrishchaka V, Senyukova O, Koepke M. Synergy of physics-based reasoning and machine learning in biomedical applications: Towards unlimited deep learning with limited data. Advances in Physics: X. 2019;4:1582361.
[12] Jena B, Saxena S, Nayak GK, Saba L, Sharma N, Suri JS. Artificial intelligence-based hybrid deep learning models for image classification: The first narrative review. Computers in Biology and Medicine. 2021;137:104803.
[13] Potla RT. Hybrid Deep Learning Models for Big Data: A Case Study in Predictive Healthcare Analytics. Distributed Learning and Broad Applications in Scientific Research. 2024;10:319-25.
[14] Gavrishchaka V, Yang Z, Miao R, Senyukova OJIJoML, Computing. Advantages of hybrid deep learning frameworks in applications with limited data. 2018;8(6):549-58.
[15] Cao C, Liu F, Tan H, Song D, Shu W, Li W, et al. Deep learning and its applications in biomedicine. 2018;16(1):17-32.
[16] Arif MS, Mukheimer A, Asif DJBD, Computing C. Enhancing the early detection of chronic kidney disease: a robust machine learning model. 2023;7(3):144.
[17] Pati A, Parhi M, Pattanayak BK, editors. An ensemble deep learning approach for Chronic kidney disease (CKD) prediction. AIP Conference Proceedings; 2023: AIP Publishing.
[18] Venkatrao K, Kareemulla SJIA. HDLNET: a hybrid deep learning network model with intelligent IoT for detection and classification of chronic kidney disease. 2023.
[19] Makino M, Yoshimoto R, Ono M, Itoko T, Katsuki T, Koseki A, et al. Artificial intelligence predicts the progression of diabetic kidney disease using big data machine learning. 2019;9(1):11862.
[20] Alsekait DM, Saleh H, Gabralla LA, Alnowaiser K, El-Sappagh S, Sahal R, et al. Toward comprehensive chronic kidney disease prediction based on ensemble deep learning models. 2023;13(6):3937.
[21] Berahmand K, Daneshfar F, Salehi ES, Li Y, Xu YJAIR. Autoencoders and their applications in machine learning: a survey. 2024;57(2):28.
[22] Li P, Pei Y, Li JJASC. A comprehensive survey on design and application of autoencoder in deep learning. 2023;138:110176.
[23] Sewak M, Sahay SK, Rathore HJJoC, Nanoscience T. An overview of deep learning architecture of deep neural networks and autoencoders. 2020;17(1):182-8.
[24] Wang W, Huang Y, Wang Y, Wang L, editors. Generalized autoencoder: A neural network framework for dimensionality reduction. Proceedings of the IEEE conference on computer vision and pattern recognition workshops; 2014.
[25] Hochreiter SJNCM-P. Long Short-term Memory. 1997.
[26] Hochreiter SJIJOU, Fuzziness, Systems K-B. Recurrent neural net learning and vanishing gradient. 1998;6(2):107-16.
[27] Noh S-HJI. Analysis of gradient vanishing of RNNs and performance comparison. 2021;12(11):442.
[28] Ioffe SJapa. Batch normalization: Accelerating deep network training by reducing internal covariate shift. 2015.
[29] Silaparasetty N. Machine learning concepts with python and the jupyter notebook environment: Using tensorflow 2.0: Springer; 2020.
[30] Akkur E, Öztürk ACJAMvUBD. Chronic Kidney Disease Prediction with Stacked Ensemble-Based Model. 2023;1(1):50-61.
[31] Lee CH, Yoon H-JJKr, practice c. Medical big data: promise and challenges. 2017;36(1):3.
[32] Sahri Z, Yusof R, Watada JJITofII. FINNIM: Iterative imputation of missing values in dissolved gas analysis dataset. 2014;10(4):2093-102.
[33] Bolikulov F, Nasimov R, Rashidov A, Akhmedov F, Young-Im CJM. Effective Methods of Categorical Data Encoding for Artificial Intelligence Algorithms. 2024;12(16):2553.
[34] Feng S, Keung J, Yu X, Xiao Y, Zhang MJI, Technology S. Investigation on the stability of SMOTE-based oversampling techniques in software defect prediction. 2021;139:106662.

[35] Kwak SK, Kim JHJKjoa. Statistical data preparation: management of missing values and outliers. 2017;70(4):407-11.
[36] Singh D, Singh BJASC. Investigating the impact of data normalization on classification performance. 2020;97:105524.
[37] Sathyanarayanan S, Tantri BRJAJBR. Confusion matrix-based performance evaluation metrics. 2024;27(4):4023-31.
[38] Muntean M, Militaru F-D, editors. Metrics for evaluating classification algorithms. Education, Research and Business Technologies: Proceedings of 21st International Conference on Informatics in Economy (IE 2022); 2023: Springer.
[39] Aldossary SMJIA, Computing S. Secured Framework for Assessment of Chronic Kidney Disease in Diabetic Patients. 2023;36(3).
[40] Babu BR, Haile MA, Haile DT, Zerihun D, Prabhakar PE, Kamyshev KJC. Prediction of Kidney Disease Using Machine Learning Algorithms. 2023(26):93-7.
[41] Venkatesan DS, Saideepika G, Sarathambekai S. A Comparative Study of Swarm-Optimized Machine Learning Models for Diabetic Kidney Disease Prediction. 2024.
[42] Shantal M, Alshareef A, Ahmid O, editors. Predicting Chronic Kidney Disease Using Filter and Wrapper Feature Selection Methods with Machine Learning Techniques. Sebha University Conference Proceedings; 2024.