International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

SASSO: Design of a Signature-Based Feature Selection Model with Ensemble Deep Learning Using Bi-LSTM and Bi-GRU for Chronic Kidney Disease Classification

[1]Anindita Khade, [2]Avaneesh Karthikeyan Iyer

Submitted: 02/05/2024 **Revised**: 15/06/2024 **Accepted**: 22/06/2024

Abstract: Chronic kidney disease (CKD) involves numerous variables, but only a few are significant for classification. The SES method, inspired by constraint-based learning in Bayesian networks, identifies essential features in CKD. Unlike traditional feature selection methods, which concentrate on a single set of features with the greatest predictive potential, the SES method can identify multiple predictive feature subsets with comparable performance. Most feature selection (FS) classifiers work better with strongly correlated data, making FS challenging in high-throughput data for finding important features and choosing the best classifier. This study is conducted on a real time data available from hospitals. This study suggests the use of the Least Absolute Shrinkage and Selection Operator (LASSO) in conjunction with the SES method, which is abbreviated as SASSO, to identify CKD features. Subsequently, a combination of Bidirectional Long Short-Term Memory (Bi-LSTM) and Bidirectional Gated Recurrent Unit (Bi-GRU) ensemble deep-learning models is proposed for CKD classification. The model's performance is measured using accuracy, precision, recall, and the F1-score. The experimental results are compared to individual classifiers, including Random Forest (RF), Naïve Bayes (NB), XGBoost (XGB), and Artificial Neural Networks (ANN).

The findings show a 6% improvement in classification accuracy with the proposed hybrid feature selection approach and the Bi-LSTM-Bi-GRU ensemble model.

Keywords*: Machine Learning, Deep Learning, Feature Selection, Bi-LSTM, Bi-GRU, SES, LASSO*

Introduction

CKD has emerged as a primary focus in contemporary medical research, according to recent studies [1,2]. CKD impacts millions globally, leading to higher morbidity, mortality, and healthcare costs[3]. Early detection and personalized treatments can significantly reduce complications and disease progression. Identifying relevant variables in CKD is crucial to enhancing classifier performance. This paper aims to develop a highly accurate CKD classification model to help doctors understand the key variables in diagnosing CKD. According to authors in [4] the existing dataset for this study available at UCI Machine Learning Repository includes critical clinical factors such as gender, age, blood pressure, laboratory results, and medical history.

Feature selection is vital in CKD classification, identifying essential features for tasks like regression and classification, retaining only the most predictive variables. It aims to improve classification performance, reduce computational cost, and eliminate unnecessary features. To tackle the challenge of numerous features in CKD diagnosis, advanced feature selection techniques are necessary to identify the most significant variables, streamlining the diagnostic process and enhancing model accuracy [5]. Effective feature selection can improve model performance through dimensionality reduction and noise elimination. Although crucial, few studies have explored effective techniques tailored to the classification task [6]. Therefore, this study develops a hybrid feature selection methodology combining the SES and LASSO feature selection methods.

Machine learning techniques have proven useful in medical diagnosis and decision-making [7]. Complex patterns can be learned and extracted from large datasets by convolutional neural networks (CNN) and recurrent neural networks (RNN), which enables precise classification in disease diagnosis and image analysis [8].

LSTM networks are useful for classification tasks because they can model and capture dependencies in sequential data. LSTMs, unlike traditional feedforward neural networks, can store information for extended periods, making them ideal for analysing sequential data like time series, text, and speech [9]. Their ability to recognize complex patterns and long-term dependencies is especially useful for classification tasks.

[1]*Assistant Professor, School of Technology Management and Engineering, NMIMS - Narsee Monjee Institute of Management Studies,Navi Mumbai, Maharashtra, India*
Email ID: aninditaac1987@gmail.com
ORCID ID: https://orcid.org/0000-0003-2616-5092
[2]*Scholar, Department of Computer Engineering, SIES Graduate School of Technology Nerul, Navi Mumbai, Maharashtra, India*
Email ID: avaneesh.karthik@gmail.com

Ensemble learning is a powerful machine learning technique that enhances classification accuracy by combining multiple classifiers. Individual classifiers' effectiveness and generalization ability are crucial to the system's success. There are two major types of ensemble learning:

- **Bagging:** Bagging improves classification accuracy by combining predictions from multiple training sets. A notable implementation of bagging is the random forest algorithm, which utilizes an ensemble of random decision trees.

- **Boosting:** Boosting is an ensemble meta-algorithm that creates classifiers by combining weak learners and iteratively training new models based on previously misclassified samples. Despite newer algorithms, Adaboost is still the most widely used boosting implementation.

Traditional machine learning methods may struggle to capture CKD's complexity fully. Therefore, ensemble deep learning techniques are necessary to identify relevant patterns in the data.

The contributions of this work are:

- Proposing a Hybrid Feature Selection method using LASSO and SES approaches for identifying CKD-related features.

- Proposing a combined Bi-LSTM and Bi-GRU ensemble deep learning model to achieve higher accuracy for CKD classification.

The study is structured as follows: Section 2 reviews the literature on CKD classification, feature selection methods, and deep learning models in healthcare. Section 3 discusses the methodology. Section 4 presents experimental findings and results and summarizes the contributions and significance of the proposed work.

Related Work

The authors of [10] investigated four effective methods for classifying CKD with SVM and AdaBoost classifiers. These methods were applied to a dataset from the UCI machine learning repository, as mentioned above. The study concluded that the AdaBoost classifier had the highest accuracy.

The authors Ahmed et al. [11] examined various machine learning techniques for early CKD prediction. Their research focused on using predictive analytics to determine how different data parameters relate to each other and the target attribute. They began with 24 characteristics along with the class property, comparing four machine learning classifiers, with the multilayer perceptron performing the best.

The authors in [12] emphasized the importance of early detection of CKD through data mining as well as machine learning to reduce patient numbers and treatment costs. They implemented ANN, SVM and KNN algorithms. They used data from the UCI repository again finding that ANN and SVM outperformed KNN in accuracy.

In [13], the authors proposed a machine learning approach to classify CKD using KNN imputation to address missing values in the UCI dataset. They tested six machine learning methods: Logistic Regression (LR), Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), naive Bayes classifier, and feedforward neural network were evaluated, with Random Forest (RF) emerging as the most effective.

The authors of [14] analyzed clinical and blood chemistry data from 680 individuals with proteinuria to predict 24-hour urinary protein outcomes using non-urine clinical indicators. The study compared five predictive models (LR, Elastic Net, lasso regression, ridge regression, SVM, and KNN) based on their precision, recall, and AU-ROC.

A Heterogeneous Modified Artificial Neural Network (HMANN) was created by Xiao et al. [15] using ultrasound images to detect, segment, and diagnose chronic renal failure early. This method reduced kidney segmentation time and improved accuracy.

Summarization of the literature review suggests that a technique of hybrid feature selection and ensemble classifiers are recommended for improving classification accuracy, particularly when dealing with nonlinear data. Keeping the above needs in mind, a hybrid selection-based ensemble classifier for CKD is proposed in this study.

Proposed Work

This work has been executed using real-time data from 1000 patients at DY Patil Hospitals in Navi Mumbai. This data comprises of 21 features. A list of features is depicted in Table 1. It is essential to assess each feature to improve the classification task. It is crucial to eradicate irrelevant features, as certain features may only make a minimal contribution to accurate classification. Most of the research has concentrated on the use of single feature selection methodologies for the classification of CKD [16,17]. Nevertheless, it is also imperative to identify features that do not contribute to the classification assignment and thereby increase the computational time in the model. To identify the critical features of CKD, a hybrid feature selection approach is necessary. Noise and model complexity can result in decreasing accuracy because of redundant features in a dataset. Overfitting may manifest when a dataset contains features that are the same or duplicate, which complicates the model's capacity to recognize patterns and relationships.

Figure 1 illustrates the proposed framework, which is divided into two phases. Phase 1 involves the task of feature selection and in Phase 2 we implement the deep learning technique of detection of CKD using ensemble learning techniques.

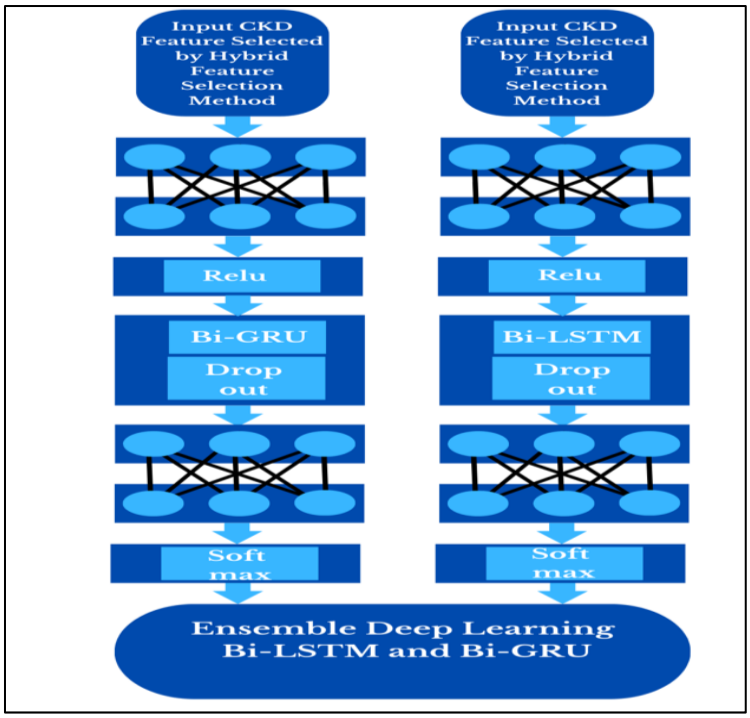


Fig 1: Proposed Framework

Table 1. CKD Feature Description

Features	**Units**	**Ranges**
age	-	25-60
gender	-	M-Male F-Female
vol	ml	2-5 ml
sg	mg/dl	0-1.25
freq	ml	0-5
sod	mg/dl	0-163
pot	mEq/L	0-47
chlo	mEq/L	0-76
phos	mEq/L	0-83
prot	gms	0-9
alb	gms	0-5
glob	gms	0-7
urea	mgs/dl	4.75-183.3
creatinine	mgs/dl	0.03-11.4
bun	mgs/dl	0-165
uricacid	mEq/L	0-391
rbc	millions/cmm	5.0-9.0
wbc	cells/cumm	6.0-11.0
pcv	cells/cumm	0-54
pe	-	yes, no
ane	-	yes, no
classification	-	ckd/notckd

SASSO feature selection

In the proposed work, the SES method is combined with LASSO regression to identify the important variable in CKD. The SES method is based on causal analysis theory and uses constraints to select features. The optimal set of predictors for a target variable, known as the Markov blanket (MB), is identified using a Bayesian network (BN) with specific assumptions [18].

Bayesian networks use a direct acyclic graph (DAG) and precise parameterization to represent multivariate distributions. Nodes represent random variables, while edges represent conditional relationships. When an edge connects two nodes in a conditional relationship, all variables between them are associated [19].

Algorithm 1, presented in pseudocode, describes the entire SASSO process. This technique calls for a dataset (D set), a variable to be targeted (tar), and two hyperparameters—a limit for variables inside a conditional set and a threshold to determine conditional independence. Through the use of these parameters, the computational complexity and resource requirements of the algorithm are constrained. Queues Q, where i ranges from 1 to N, contain K variables that are equivalent to one another [20].

All variables v added to Set s are part of an empty set created during initialization.

The list of variables is represented by L → v, and each variable is only equivalent to Qi → i. The algorithm loops and selects the variable with the highest correlation with Tar based on a subset of selected variables. Variables that are not correlated with Tar are excluded.

Excluded variables *v* are ineligible for addition to *Set s*. Before eliminating *v* from *Set s*, the method searches for a variable *p* in *subset* that matches *v* by validating when $sub\ set \rightarrow sub\ set * \{v\} \backslash \{p\}$. The LASSO model receives these selected features. LASSO feature selection uses ridge regression and subset selection to improve model interpretation and prediction[21]. It maintains one feature correlation while minimizing others to zero. All regression methods aim to determine coefficients for existing features to ensure that each data sample yields the desired result. The LASSO method truncates coefficients to zero by adjusting them by a constant value λ. Reducing the residual sum of squares lowers coefficient absolute values.

Algorithm 1

1. Input: CKD Datasets.
2. Definitions:
- D: Represents datasets.
- v: Independent variables.
- tar: Represents dataset target variables.
- threshold: Represents threshold.
3. Initialize:
- A collection of data on n independent predicting variables i.
- The variable t is the target.
- Indicator of threshold th variable.
- Max: Set representing the maximum variable.
- A maximum of N variables (Q_i=1 to N) can be used to generate a signature by selecting only one from each set.
- List of variables L←v.
- Variables selected.
- Set s←∅.
- Equivalence sets.
- Q_i←i, for i=1 to N.
4. Algorithm:
   While L!=0
   For all v belonging to {L\Set s} do:
   L← L\{v}
   Set s ←Set s\{v}
   Remove irrelevant features
5. Input CKD features to LASSO feature selection method
6.
7. List the selected features from SASSO

Ensemble Deep Learning with Bi-LSTM and Bi-GRU

In phase 2, we proposed an ensemble-based deep learning approach for classifying CKD. While previous work predominantly utilized individual classifiers for this task, CKD datasets are inherently nonlinear. Using individual classifiers may not achieve high accuracy due to the dataset's nonlinear nature. Hence, our approach adopted ensemble-based deep learning to address these challenges[22].

Figure 1 depicts the proposed ensemble learning architecture. This architecture involves training two deep learning models on the same dataset and combining their predictions to produce a final classification result. Each model has a binary output for chronic kidney disease (class 1) or absence of chronic kidney disease (class 0). Pilot tests demonstrated that LSTM and GRU models outperformed other models, especially for class 1 (chronic kidney disease) and class 0 (absence) [23].

Bi-LSTM neural networks were developed using a series of interconnected neural network modules to address the problem of long-term dependency. A Bi-LSTM-based model is constructed to depict the development of CKD because of LSTM's ability to understand long-term dependencies In order to predict CKD, we have taken into account three layers: the Pre-Fully Connected Layer, the Cells Layer, and the Post-Fully Connected Layer, as illustrated in Figure 1. The First Fully Connected layer has a ReLU function and one fully connected layer; the

Second Layer has one Long Short-Term Memory (LSTM) layer and a Dropout Wrapper; and the Third Layer has one completely connected layer and a softmax layer. Bi-LSTM is a significantly more potent instrument than the standard or unidirectional LSTM due to its ability to incorporate information from both past and future timestamps, thereby generating credible reverse evaluation figures that also facilitate cross-validation of the results[24]. Bi-LSTM can enhance the prediction of a machine learning model by training it with values from both the timestamps and the learning rate, thereby increasing the degree of substantiation[25].Figure 2 describes the architecture of Bi-LSTM.

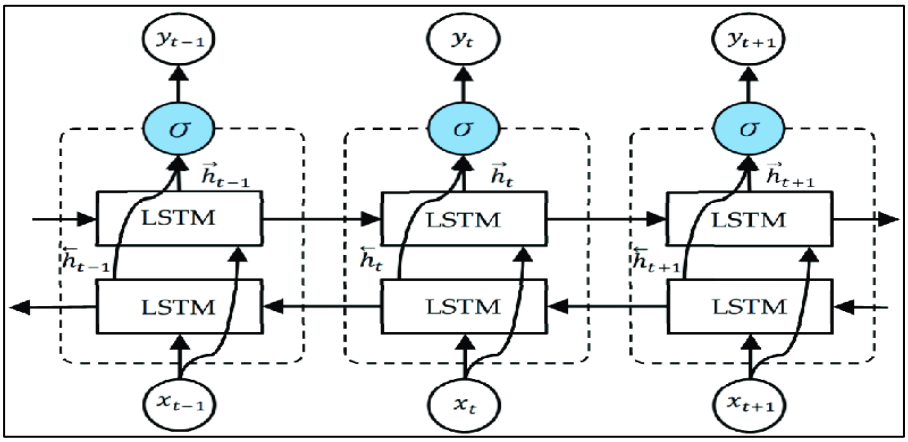


Fig 2: Bi-LSTM Architecture

In contrast to Bi-LSTM, Bi-GRU uses fewer gates in its RNN architecture. In a Bi-GRU cell unit, a single gate controls both the input and forget functions[26].Figure 3 describes the architecture of Bi-GRU.

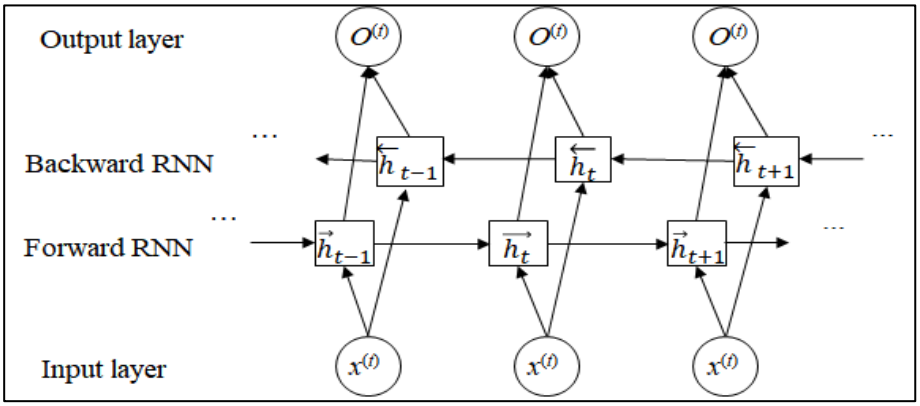


Fig 2: Bi-GRU Architecture

Artificial Neural Networks (ANN)

"Artificial Neural Network" describes a subfield of AI that draws design cues from how the brain really works. To simulate the way the brain's own neural networks, operate, computer scientists have developed what is known as an artificial neural network (ANN) [27]. The building blocks of an artificial neural network are neurons linked together in a hierarchical structure, much like the human brain. It is usual practice to refer to these neural structures as nodes. A lot of layers make it up.

The programmer determines the format that this layer may take inputs in.

The input and output layers are separated by the concealed layer. It finds hidden patterns and traits by doing all the necessary calculations.

Transmitted via this layer is the final output, which is the product of the input's modification by the hidden layer.

The ANN takes in data and then computes the weighted sum of that data, which incorporates a bias factor. The activation function is given the weighted total to produce the output [28]. For each node, the activation function is what defines how it fires. A number of activation functions are available, each tailored to a different task.

In Equation 1, the formula for ANN is described.

$$\sum_{i=1}^{n} w_i * x_i + b \quad (1)$$

XG Boost (XGB)

Boosting is an ensemble learning technique that aims to minimize training errors by combining poor learners to produce stronger ones. A model is fitted using a random sample of data, and it is subsequently trained

progressively. This is known as the enhancing technique. The constraints of its predecessor are pushed to be surpassed by each subsequent model[29].

The weak rules of each classifier are combined to generate a single strong prediction rule during each iteration. The primary objective of the XGBoost algorithm is to enhance the efficacy and performance of computational processes. The sequential analysis of the dataset by the Gradient Descent Boosting algorithm results in delayed output computation and a reduction in computational efficiency. Consequently, the model's efficacy is substantially enhanced through the implementation of XGBoost.

Naïve Bayes (NB)

Naive Bayes classifiers are a collection of classification algorithms based on Bayes' Theorem. It is a collection of algorithms that share a common principle: each pair of features being classified is independent of one another. Initially, we will look at a dataset.

The Naïve Bayes classifier is a simple yet effective classification algorithm. It is essential for the rapid development of machine learning models capable of making quick predictions[31].

The Naïve Bayes algorithm is used for solving classification problems. It is widely used in the classification of texts. Text classification tasks involve high-dimensional data, with each word representing a feature. It is used in applications such as sentiment detection, spam filtering, and rating classification[32]. The speed of Naive Bayes is a significant advantage. Because of the high dimension of the data, it is quick and easy to make predictions[33].

Random forest classifier

The random forest algorithm is an ensemble learning approach that is founded on decision tree learning. A single classifier may be unable to accurately classify test data due to its inability to differentiate between noise and patterns based on sample data, as per ensemble learning. The provided data sets are used to train multiple decision trees (n trees) using sampling with replacement in the random forest classifier. A random set of three attributes is used to train each tree. Following the data evaluation, the ultimate output is determined by the consensus decision of the n trees[33].

Experimental Results

In the experimental work, we considered CKD datasets with 21 features. The final variable is the class variable, which has two options: *CKD Yes* and *CKD No*. We considered 1000 samples, 650 of which were classified as *CKD Yes* and 350 as *CKD No*.

Table 2 shows the best features selected using the SASSO algorithm. These are the top 13 features that can be used in decision-making processes.

Table 2: Features and Rankings selected from SASSO

SASSO	
Feature	**Ranking**
sc	0.629
bu	0.611
sod	0.588
alb	0.543
phos	0.449
prot	0.439
sg	0.401
glob	0.309
bun	0.277
wbc	0.232
chlo	0.132
rbc	0.086
pcv	0.074

Table 3 shows the results of comparison before and after implementing the hybrid feature selection algorithm.

Accuracy, precision, recall, and F1 score were some of the performance metrics used to compare the suggested method to various state-of-the-art algorithms.

The amount of times a model's predictions turn out to be right is called its accuracy. (2) The formula for accuracy is (TP+TN) divided by (TP+TN+FP+FN).

The number of positive classifications that are calculated to be accurate is known as precision.

The formula for precision is TP divided by the sum of TP and FP.

The number of positive classifications that the classifier properly detected is called recall.

Formula: Recall equals TP divided by the sum of TP and FN (4)

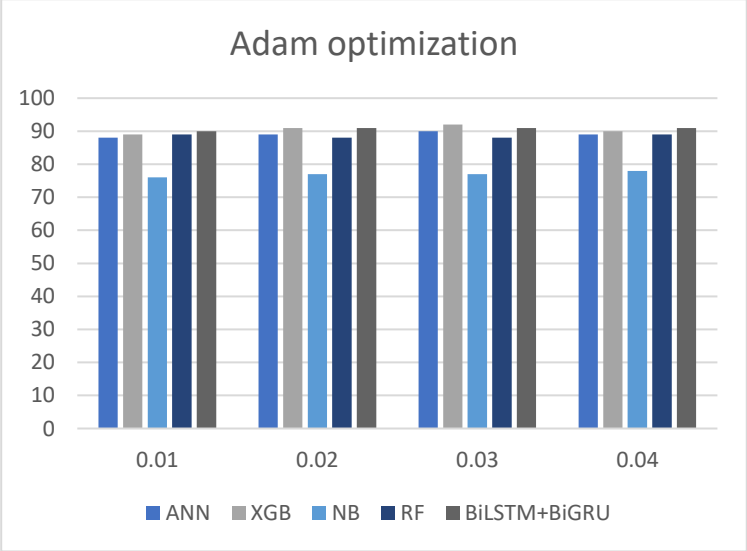
The F1 Score measures the frequency with which a model has produced accurate predictions throughout the whole dataset.

The F1 score can be calculated as (2*Precision*Recall)/(Precision Recall). (5)

Table 3. Results Comparison

Model	Feature Selection	Accuracy	Precision	Recall	F-score
ANN	No	0.87	0.86	0.84	0.83
XGB	No	0.88	0.83	0.86	0.85
NB	No	0.76	0.74	0.73	0.75
RF	No	0.88	0.85	0.86	0.86
Bi-LSTM+Bi-GRU	No	0.90	0.91	0.91	0.91
ANN	SASSO	0.90	0.90	0.90	0.91
XGB	SASSO	0.92	0.92	0.92	0.91
NB	SASSO	0.89	0.85	0.84	0.86
RF	SASSO	0.94	0.93	0.93	0.94
Bi-LSTM+Bi-GRU	SASSO	0.97	0.97	0.93	0.94

The hybrid feature selection method is critical as it identifies the most pertinent variables in a dataset, effectively reducing dimensionality while retaining key information. In our experiments, we tested several learning rates—0.01, 0.02, 0.03, and 0.04—across different optimizers like Adam and Stochastic Gradient Descent to optimize our models. After evaluating all combinations of learning rates and optimizers, we determined that Stochastic Gradient Descent (SGD) paired with a learning rate of 0.02 yielded the highest accuracy. This optimal configuration is detailed in Figures 4 and 5.



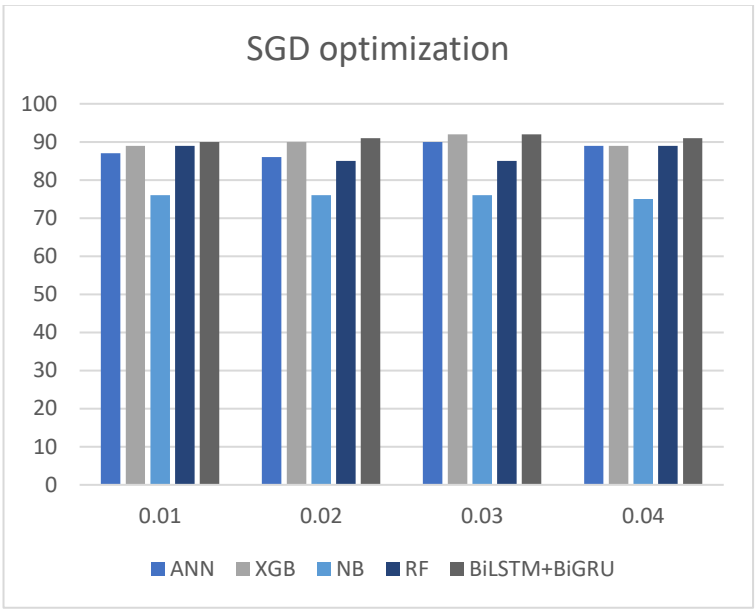


Fig 4: Without Feature Selection

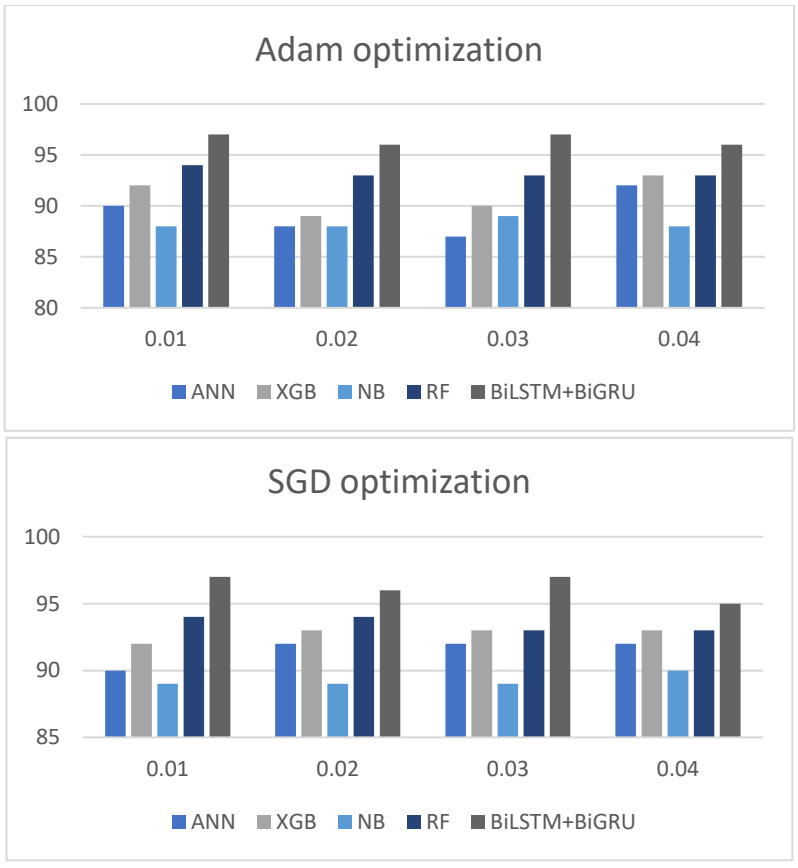


Fig 5: With SASSO Feature Selection

In our experimental study, metrics such as accuracy, precision, recall, and F1 score were utilized in our evaluation of the suggested model. These metrics are specified in Equations 2, 3, 4, and 5.

Comparative results with ANN, XGB, NB, and Random Forest methods, both with and without feature selection, are summarized in Table 3. Our methodology exhibited significant performance gains, surpassing alternative methods by a margin of 6%. This improvement is attributed to the ensemble model leveraging Bi-LSTM and Bi-GRU, which excel in learning and capturing long-term dependencies in sequential data. The combination of Bi-LSTM and Bi-GRU enhances generalization capabilities, thereby improving performance on unseen data.

Future research directions include exploring advanced deep learning architectures like autoencoders to further enhance classification performance. Additionally, conducting comparative analyses with state-of-the-art

feature selection and classification methods would provide valuable insights into the strengths and weaknesses of our approach. Additional enhancements to prediction skills could be achieved through the incorporation of multimodal data sources, such as genetic information or images obtained from histology. In conclusion, it would be essential to conduct a clinical validation research in order to evaluate the applicability and performance of our suggested model in relation to healthcare settings in the real world.

Conclusion

The proposed project is divided into two phases. In the initial phase, a composite feature selection method was employed to identify highly predictive features in for detection of CKD. During the subsequent phase, an ensemble deep learning method based on ensembles was suggested for the classification of CKD. The ensemble-based deep learning method that has been proposed is a hybrid of Bi-LSTM and Bi-GRU. We used a variety of learning rates in our experimental work to fine-tune the models: 0.01, 0.02, 0.03, and 0.04. In the proposed model, we evaluated the optimization algorithms Adam and Stochastic Gradient Descent. In comparison to other learning rates and optimizers, our experiments demonstrated that the SGD optimizer, in conjunction with a learning rate of 0.02, yielded the highest accuracy. The experimental results are compared to those of individual classifiers, including ANN, RF, XGB and Naïve Bayes. The results suggest a 6% increase in classification accuracy when the proposed hybrid feature selection approach and ensemble Deep Learning model Bi-LSTM and Bi-GRU are taken into account. Adjusting the hyperparameters of deep learning models, including the number of layers, hidden units, learning rates, and dropout rates, could potentially enhance the model's performance in the future. Bayesian optimization or genetic algorithms are automated methods that can be employed to efficiently investigate the hyperparameter space and identify the optimal configurations.

Competing Interests

The authors declare no conflict of interest.

Funding

The authors received no funding for this work.

Data Availability

The data will be available on request since it contains patient information

References

[1] Khade, A. A., Vidhate, A. V., & Vidhate, D. A Comparative Analysis of Applied AI Techniques for an Early Prediction of Chronic Kidney Disease. Proceedings - 2nd International Conference on Smart Electronics and Communication, ICOSEC 2021; 1386–1392. https://doi.org/10.1109/ICOSEC51865.2021.9591869

[2] Chittora, P., Chaurasia, S., Chakrabarti, P., Kumawat, G., Chakrabarti, T., Leonowicz, Z., Jasinski,M., Jasinski, L., Gono, R., Jasinska, E., & Bolshev, V. Prediction of Chronic Kidney Disease - A Machine Learning Perspective. IEEE Access, 9(January),2021; 17312–17334. https://doi.org/10.1109/ACCESS.2021.3053763

[3] Dare, A. J., Fu, S. H., Patra, J., Rodriguez, P. S., Thakur, J. S., & Jha, P. Renal failure deathsand their risk factors in India 2001–13: nationally representative estimates from the Million Death Study. The Lancet Global Health, 5(1), 2017; e89–e95. https://doi.org/10.1016/S2214-109X(16)30308-4

[4] Ifraz, G. M., Rashid, M. H., Tazin, T., Bourouis, S., & Khan, M. M. Comparative Analysisfor Prediction of Kidney Disease Using Intelligent Machine Learning Methods. Computational and Mathematical Methods in Medicine, 2021; https://doi.org/10.1155/2021/6141470

[5] Yuan, Q., Zhang, H., Deng, T., Tang, S., Yuan, X., Tang, W., Xie, Y., Ge, H., Wang, X., Zhou, Q.,& Xiao, X.. Role of artificial intelligence in kidney disease. International Journal of Medical Sciences, 17(7),2020; 970–984. https://doi.org/10.7150/ijms.42078

[6] Bhaskar, N., & Suchetha, M. A Computationally Efficient Correlational Neural Network for Automated Prediction of Chronic Kidney Disease. IRBM, 42(4), 2020;268–276. https://doi.org/10.1016/j.irbm.2020.07.002

[7] Muslim, M. A., Kurniawati, I., & Sugiharti, E. Expert system diagnosis chronic kidney disease based on Mamdani fuzzy inference system. Journal of Theoretical and Applied Information Technology,2015; 78(1), 70–75.

[8] [8] Hu, D., Nie, F., & Li, X. Deep linear discriminant analysis hashing. Scientia Sinica Information is 51(2), 2021;279–293. https://doi.org/10.1360/SSI-2019-0175

[9] [9] Singh, V., Asari, V. K., & Rajasekaran, R. A Deep Neural Network for Early Detection and Prediction of Chronic Kidney Disease. Diagnostics, 2022; 12(1),1–22.https://doi.org/10.3390/diagnostics12010116

[10] Almansour, N. A., Syed, H. F., Khayat, N. R., Altheeb, R. K., Juri, R. E., Alhiyafi, J., Alrashed, S., & Olatunji, S. O. Neural network and support vector machine for the prediction of chronic kidney disease: A comparative study. Computers in Biology and Medicine, 2019; 109(April), 101–111.

https://doi.org/10.1016/j.compbiomed.2019.04.017

[11] Senan, E. M., Al-Adhaileh, M. H., Alsaade, et al. Diagnosis of Chronic Kidney Disease Using Effective Classification Algorithms and Recursive Feature Elimination Techniques. Journal of Healthcare Engineering, 2021;. https://doi.org/10.1155/2021/1004767

[12] Rady, E. H. A., & Anwar, A. S. Prediction of kidney disease stages using data mining algorithms. Informatics in Medicine Unlocked, 15(December 2018), 2019; 100178. https://doi.org/10.1016/j.imu.2019.100178

[13] Khamparia, A., Saini, G., Pandey, B., Tiwari, S., Gupta, D., & Khanna, A. KDSAE: Chronic kidney disease classification with multimedia data learning using deep stacked autoencoder network. Multimedia Tools and Applications, 79(47–48), 2020; 35425–35440. https://doi.org/10.1007/s11042-019-07839-z

[14] Song, X., Waitman, L. R., Yu, A. S. L., et al, Longitudinal risk prediction of chronic kidney disease in diabetic patients using a temporal-enhanced gradient boosting machine: Retrospective cohort study. JMIR Medical Informatics, 2020;8(1), 1–16. https://doi.org/10.2196/15510

[15] Pasadana, I. A., Hartama, D., Zarlis, M., et al. Chronic Kidney Disease Prediction by Using Different Decision Tree Techniques. Journal of Physics: Conference Series,2019; 1255(1). https://doi.org/10.1088/1742-6596/1255/1/012024

[16] Krishnamurthy, S., Kapeleshh, K. S., Dovgan, E., et al. Machine learning prediction models for chronic kidney disease using national health insurance claim data in Taiwan. Healthcare (Switzerland), 2021;9(5), 1–13. https://doi.org/10.3390/healthcare9050546

[17] Neves, J., Martins, M. R., Vilhena, J., Neves, J., Gomes, S., Abelha, A., Machado, J., & Vicente,

[18] H. A Soft Computing Approach to Kidney Diseases Evaluation. Journal of MedicalSystems, 2015;39(10). https://doi.org/10.1007/s10916-015-0313-4

[19] Fokas, A. S., Dikaios, N., & Kastis, G. A. Mathematical models and deep learning for predicting the number of individuals reported to be infected with SARS-CoV-2. Journal of the Royal Society Interface,2020; 17(169). https://doi.org/10.1098/rsif.2020.0494

[20] Chen, X., & Jeong, J. C. Enhanced Recursive Feature Elimination. January 2008.2014; https://doi.org/10.1109/ICMLA.2007.35

[21] Shastri, S., Kour, P., Kumar, S., Singh, K., & Mansotra, V. XGBoost: A novel Grading- Ada Boostensemble approach for automatic identification of erythemato-squamous disease. International Journal of Information Technology (Singapore), 2021; 13(3), 959–971. https://doi.org/10.1007/s41870-020-00589-4

[22] Khanday, A. M. U. D., Rabani, S. T., Khan, Q. R., Rouf, N., & Mohi Ud Din, M. Machine learning-based approaches for detecting COVID-19 using clinical text data. International Journal of Information Technology (Singapore), 2020;12(3), 731–739. https://doi.org/10.1007/s41870-020-00495-9

[23] Dasari, S. K., & Prasad, V. A novel and proposed comprehensive methodology using deep convolutional neural networks for flue-cured tobacco leaves classification. InternationalJournal of Information Technology (Singapore), 2019;11(1), 107–117. https://doi.org/10.1007/s41870-018-0174-4

[24] [23] Khade, A. A. and Vidhate, A. V. Application of Artificial Intelligence Techniques in the Early-Stage Detection of Chronic Kidney Disease, Data Science-Techniques and Intelligent Applications(NewYork), 2022.

[25] H.-C. Lee et al., "Prediction of Acute Kidney Injury after Liver Transplantation: Machine Learning Approaches vs. Logistic Regression Model," Journal of Clinical Medicine, vol. 7, no. 11, p. 428, Nov. 2018, doi: 10.3390/jcm7110428

[26] Elhoseny, M.; Shankar, K.; Uthayakumar, J. Intelligent diagnostic prediction and classification system for chronic kidney disease. Sci. Rep. 2019, 9, 9583. [CrossRef] [PubMed].

[27] Vasquez-Morales, G.R.; Martinez-Monterrubio, S.M.; Moreno-Ger, P.; Recio-Garcia, J.A. Explainable Prediction of Chronic Renal Disease in the Colombian Population Using Neural Networks and Case-Based Reasoning. IEEE Access 2019, 7, 152900–152910.

[28] Senan, E.M.; Al-Adhaileh, M.H.; Alsaade, F.W.; Aldhyani, T.H.H.; Alqarni, A.A.; Alsharif, N.; Uddin, M.I.; Alahmadi, A.H.;

[29] Jadhav, M.E.; Alzahrani, M.Y. Diagnosis of Chronic Kidney Disease Using Effective Classification Algorithms and Recursive

[30] Feature Elimination Techniques. J. Healthc. Eng. 2021, 2021, 1004767.

[31] Krishnamurthy, S.; Ks, K.; Dovgan, E.; Luštrek, M.; Piletiˇc, B.G.; Srinivasan, K.; Li, Y.-C.; Gradišek, A.; Syed-Abdul, S. Machine learning prediction models for chronic kidney disease using national health insurance claim data in Taiwan. Healthcare 2021, 9, 546.

[32] Polat, H.; Mehr, H.D.; Cetin, A. Diagnosis of chronic kidney disease based on support vector machine by feature selection methods. J. Med. Syst. 2020, 41, 55..

[33] S. K. Agarwal, S. C. Dash, M. Irshad, S. Raju, R.

Singh, and R. M. Pandey, “Prevalence of chronic renal failure in adults in Delhi, India,” Nephrology Dialysis Transplantation, vol. 20, no. 8, pp. 1638–1642, Apr. 2005, doi: 10.1093/ndt/gfh855.

[34] S. Singh, S. Shreevastava, T. Som, and G. Somani, “A fuzzy similarity-based rough set approach for attribute selection in set valued information systems,” Soft Computing, vol. 24, no. 6, pp. 4675–4691, Jul. 2019, doi: 10.1007/s00500-019-04228-4.

[35] Khade, A., Vidhate, A.V. & Vidhate, D. FFN-XGB-design of a hybrid feed forward neural network and extreme gradient boosting model for early prediction of chronic kidney disease. Int J Syst Assur Eng Manag 2023. https://doi.org/10.1007/s13198-023-01993-2

[36] Khade, A.., Vidhate, A. V., & Vidhate, D., Design of an Optimized Self-Acclimation Graded Boolean PSO with Back Propagation Model and Cuckoo Search Heuristics for Automatic Prediction of Chronic Kidney Disease. Journal of Mobile Multimedia, 19(06), 1395–1414. https://doi.org/10.13052/jmm1550-4646.1962