ISSN 2277-2685
IJESR/October-December 2024/ Volume-14/Issue-4/162-183
Surendar Rama Sitaraman *et. al.*, / International Journal of Engineering & Science Research

BI-DIRECTIONAL LSTM WITH REGRESSIVE DROPOUT AND GENERIC FUZZY LOGIC ALONG WITH FEDERATED LEARNING AND EDGE AI-ENABLED IOHT FOR PREDICTING CHRONIC KIDNEY DISEASE

Surendar Rama Sitaraman, Intel Corporation, California, USA
surendar.rama.sitaraman@ieee.org
Poovendran Alagarsundaram, Humetis Technologies Inc, Kingston, NJ, USA poovendrana@ieee.org
Harikumar Nagarajan, Global Data Mart Inc (GDM), New Jersey, USA harikumarnagarajan@ieee.org
Venkata Surya Bhavana Harish Gollavilli, Under Armour, Maryland, USA venkataharish@ieee.org
Kalyan Gattupalli, Yash Tek inc, Ontario, Canada
kalyangattupalli@ieee.org
Dr.S.Jayanthi M.E.,Ph.D., Principal, Tagore Institute of Engineering and Technology Deviyakurichi, Thalaivasal (TK), Salem
sjayanthi@ieee.org

ABSTRACT: *Chronic Kidney Disease (CKD): A growing public health concern Appropriate and timely diagnosis of chronic diseases is important in determining subsequent interventions. In this research, we present a CKD prediction model with FL, edge AI and Bi-LSTM along with Regressive Dropout and GELU activation to boost the performance of CKD prediction. In addition, Generic Fuzzy Logic (G-Fuzzy) increases the accuracy of CKD stage classification. The computational complexity of the proposed model is further accelerated with a Granular Information-based Krill Herd Algorithm (GI-KHA) that performs feature selection. Data Results: Achieves better prediction ability (98.96%) than traditional approaches The approach is scalable, respects patient data privacy and thus applicable in the healthcare settings where real-time diagnosis of CKD could be leveraged.*

Objective: *Creating a CKD detection model by integrating Edge AI with Federated Learning, and improving the accuracy rate, privacy, as well as real-time usage.*

Methods: *This model uses Bi-LSTM with GELU activation, Regressive Dropout, G-Fuzzy Logic for CKD classification and GI-KHA for feature selection. Privacy in a Distributed Data training is achieved by using Federated Learning.*

Results: *The sensitivity, precision and recall were superior to the traditional model with an accuracy of 98.96%.*

Conclusion: *Our system presents a method for early detection of CKD which is efficient, scalable and privacy-preserving, whereby the proposed solution has the potential to provide enhancements in patient care.*

Keywords: *Chronic Kidney Disease (CKD), Federated Learning, Edge AI, Bi-LSTM, GELU, Regressive Dropout, Generic Fuzzy Logic, Krill Herd Algorithm, Healthcare Prediction.*

1. INTRODUCTION

If Chronic Kidney Disease (CKD) develops, the fact that it is often insidious and difficult to diagnose means you run a high risk of ending up with kidney failure if not detected early enough. Integration of Federated Learning (FL) and Edge AI with Bi-Directional Long Short-Term Memory (Bi-LSTM) networks has shown promising results in CKD prediction when integrated over Internet of Health Things (IoHT) *(Xu et al., 2020)*. The GELU (Generalised Exponential Linear Unit) Bi-LSTM with Regressive Dropout and Generic Fuzzy Logic system are some of the advanced methods used in predicting CKD, by *(Janani et al., 2022)*. Although Edge AI enables real-time processing—necessary for quick diagnosis and intervention—the proposed model uses FL to maintain data privacy as well as enhance the overall performance of it *(Mosaiyebzadeh et al., 2023)*. They argued that the model enhances prediction accuracy via Bi-LSTM capturing temporal dependencies in patient data efficiently *(Galles et al., 2023)*. Another set of studies introduce Phase Mapping by the Generic Fuzzy Logic system in order to provide a detailed aspect on diagnostics based on pertaminating CKD, breaking down the prediction process even further *(Iroegbu et al., 2021; Praveen et al., 2022)*. The inclusion of GELU will help to make the model more robust for reduced risks in overfitting *(Jawad et al., 2024; Saif et al., 2024)*. This method, through early and accurate diagnosis not only advances the IoHT field but also provides a scalable solution for CKD prediction which may result in better outcomes. In Surendar Rama Sitaraman's (2023) investigation of AI integration in healthcare, Turkey's National AI Strategy is highlighted. AI improves patient outcomes, makes the most use of available resources, and promotes customized care, making Turkey a pioneer in the efficient and innovative use of AI in healthcare.

Chronic Kidney Disease (CKD) is a serious public health problem due to asymptomaticity in the initial stages and may end up with kidney failure if not treated early. Until now, conventional clinical examinations such as urine assays and glomerular filtration rate (GFR) measures had limited utility in the diagnosis of illness that often presented later on *(Kitzler and Chun, 2023)*. That has become a thing of the past due to advent in AI and machine learning in healthcare, such that earlier diagnosis leading to better clinical overview on CKD will be made possible *(Durga and Sudhakar, 2023)*. The integration of FL with Edge AI for the prediction of CKD is a huge technology leap, especially in healthcare *(Badidi 2023; Reddy 2024)*. PrivacyFabric: FL exploits the concept of federated micro-aggregations to enable decentralized data processing while still being able to retain all advantages that come along with collaborative learning and keep even more sensitive patient health related data on-device *(Sadilek et al., 2021)*. When it is related to health and in a hospital atmosphere, privacy of the patient should be one of your priorities. To improve intelligent medical diagnostics, Surendar Rama Sitaraman (2021) introduces Crowd Search Optimisation (CSO), a metaheuristic algorithm modeled after crows' foraging behavior. CSO outperforms conventional techniques in CNN and LSTM hyperparameter optimization for disease diagnosis, resulting in improved model performance and accuracy.

The most significant technical jump in healthcare, especially CKD prediction, is FL & Edge AI integration. Maintain Collaborative learning with FL - one of the big assets in enabling even more decentralised data processing as possible. Preserving Patient privacy by localizing medical data. This is extremely important in healthcare settings for obvious reasons tied to ensuring patient privacy. Edge AI solves this problem by reducing latency and improving system responsiveness through real-time data analysis at the point of treatment *(Chen et al., 2022)*. Proposed model: because Bi-Directional LSTM (Bi-LSTM) models are excellent at capturing the time-line characteristic of medical data, this results in a higher prediction accuracy so its application is significant *(Hoang et al., 2024)*. The correcting overfitting and lack of generalisation between diverse patient datasets broadened the role of Generalised Exponential Linear Unit (GELU) used in Bi-LSTM framework as well as Regressive Dropout hence enhancing performance by a great amount *(Majid, 2023; Gopi et al., 2023)*. The combination of FL, Edge AI and advanced neural network approaches represents an innovative way in CKD prediction *(Budrionis et al., 2021; Fonseca, 2023)*. The use of AI and Big Data Analytics in m-Health technologies to improve healthcare delivery is highlighted by *Surendar Rama Sitaraman (2020),* whose neural networks achieve 92% accuracy. Despite encouraging developments, there are still difficulties in managing unstructured data from wearables and protecting data privacy.

1.1. Problem Statement

Saif et al. This work (2024) tackles the understudied problem of early Chronic Kidney Disease (CKD) prediction and introduces a framework that provides a deep learning based predictive model using a more advanced ensemble learning approach to predict CKD on a set within a specific time range. The project humanises this contribution by focusing on the models (OM1-4) i.e. CNN, LSTM and Transformer as well as comparing analyzed with legacy solutions only aimed at improving detection of emotions currently used to face selling biases or proactive interventions for 6-to12 month forecasts coming close to bridging existing gaps between both aspects Compared to previous models, this work focuses on solving problems like data imbalance and feature selection and algorithm parameter optimization to improvedly accurate prediction of CKD at a very early stage.

1.2. Objectives

- For developing a CKD prediction model while protecting patient privacy by utilising Edge AI and Federated Learning.
- Regressive Dropout and Bi-Directional LSTM with GELU for better predictions.
- For implementing generic fuzzy logic into the CKD stage categorisation process.
- Improve early intervention with real-time CKD prediction by Edge AI
- Develop a model to predict CKD with high accuracy and generalisability across different settings of care.

2. RELATED WORK

Majid (2023) applies new methodologies such as advanced data mining techniques and Ensemble learning to enhance the diagnosis of Chronic Kidney Disease (CKD). This study combined multiple machine learning models which significantly improved the diagnostic accuracy and consistency in contrast to using single models. Sophisticated data mining techniques are employed to uncover major patterns in medical datasets. We model the data with these patterns to train all the ensemble models. This approach would significantly enhance the detection of CKD--better equipping healthcare professionals to proactively manage this disease and hopefully lead better patient outcomes in clinical care.

Surendar Rama Sitaraman (2022) emphasizes how AI is revolutionizing radiology, especially with CNNs for image analysis automation and VAEs for data augmentation. AI promises better health outcomes and diagnostic accuracy despite issues like data privacy and interpretability of models.

Zisser and Aran (2024) generated a transformer-based model for predicting the time-to-event of chronic kidney disease (CKD) progression with much greater accuracy than Cox proportional hazards or other more traditional models. The model provides robustness and predictive power compared to better performing models due to increased clinical decision-making by patient longitudinal data used, an important advance for CKD patient-centered healthcare, this strategy is validated across multiple cohorts of patients with kidney disease It permits earlier and more targeted therapies.

Gopi et al. (2023), IoT enables early detection of Chronic Kidney Disease (CKD) in the medical sector using a Temporal Convolutional Network (TCN). The patient data is collected by IoT devices and then analysed into a software application which can be accessed by medical professionals. So, it is very common and someone with chronic kidney disease (CKD) can have disastrous consequences if not identified in early stages. In this study, we built a TCN-based classifier for the CKD dataset from Kaggle applying a deep learning model and Feature extraction with Latent Dirichlet Allocation LDA features. The TCN method apparently yields accurate predictions from CKD stage, which could help in identifying the early stages of progression enhancing better patient outcomes this is important.

Huang et al. (2023) created a predictive model for acute kidney damage (AKI) in critically ill patients across several Taiwanese hospitals by utilizing federated learning (FL). The study used extreme gradient boosting, neural networks, and random forests to train and test AKI prediction models by combining data from Taichung Veterans General Hospital and four other referral hospitals. With both extensive and condensed feature sets, the models' accuracy was comparable, as indicated by their AUROC curves, which hovered around 0.90. Model performance at the participating centers was marginally improved by FL. This method improved prediction across various healthcare settings by accurately predicting AKI with a 24-hour lead time.

Categories and non-categorical attributes were preprocessed with machine learning algorithms to develop a model for early detection of Chronic Kidney Disease (CKD) by Pal (2023). This is an attempt to establish high performance baseline classifiers on category and non-categorical variables independently, which are then fused using a majority vote approach. The combination model increased the total accuracy by 3% on what it was without it This approach helps with early identification of CKD for the public, and assists physicians in devising their treatment plans so that its classification abilities are overall improved.

Hoang et al., in a study of patients with stages 3–5 chronic kidney disease (CKD), sought to predict the timing at which they would commence maintenance dialysis. The remaining study (Afolabi 2024) constructed and validated a machine learning (ML) model. The study from Taipei Medical University developed and assessed predictive models based on patient demographics, comorbidities, lab results, along with medication information (from parallel testing) using historical data. The artificial neural network showed high performance and achieved AUC values from 0.96 for prediction at year-1 to 0.92 prediction on year-3 (Table in supplementary). The important characteristics for albuminuria and the baseline estimated glomerular filtration rate were also included. This approach to machine learning, which is suited for personalized predictions based on AI-PHI data in PRO consisting of a large number of predictors with small sample size at individual level can aid clinical decision and may lead to improved patient outcomes.

Ashafuddula et al. (2023) dealt with common problems in clinical data like missing values, noise and redundancy to construct an intelligent diagnosis system for the detection of early-stage Chronic Kidney Disease (CKD). Therefore, for better performance of the model they incorporate in their fully automated machine learning strategy data balance, Feature Space Reduction(FSR) and feature selection (FS). An ensemble of adaptive boosting, a logistic regression and passive-aggressive classifiers achieved an accuracy rate as high as 96.48% in case prediction CKD using real-world patients data collected from Bangladesh. The approach appears to be a very useful clinical diagnostic and timely intervention strategy since it reduces the prediction time significantly leading to early detection.

Akinyemi et al. (2023) investigated the possibility that machine learning (ML) and better data processing might predict chronic kidney disease (CKD). CKD is not just a global health issue but one with significant implications that advances in phase detection are essential. There are multiple Machine Learning (ML) techniques like Decision Tree(DT), Naïve Bayes(NB), and Random Forest, used in this study for analyzing different variables including Age, Albumin levels etc. Efficient data preprocessing improves predictive accuracy. Random Forest had the best accuracy (96.25%), and XGBoost/NB were perfect in sensitivity. The importance of needing to accurately select a model for prediction as well as handling missing data has been under-listed by this study.

Saif et al. (2024) presented a new methodology of utilizing an ensemble of deep learning models and optimizer for the early prediction model for Chronic Kidney Diseases (CKD). The method combines multiple deep learning models, in such a way that not only the prediction accuracy is higher (better) but also complex patterns can be discovered between CKD data and other factors. Unlike conventional techniques, it includes optimisers within the model for enhancing performance and hence providing more accuracy. This study represents a significant advance in the diagnosis of CKD by improving early detection and treatment.

Surendar Rama Sitaraman (2022) investigates the way edge computing, using methods like federated learning and homomorphic encryption, might improve IoT security and privacy through anonymized AI. The study demonstrates that confidential data is successfully protected by anonymized AI without sacrificing functionality.

3. METHODOLOGY

This section describes the exact technique of the proposed system. This architecture uses state-of-the-art machine learning methods Bi-Directional Long Short-Term Memory (Bi-LSTM) with GELU activation and Regressive Dropout for enhancing CKD prediction, along Federated Learning (FL)-based decentralised data processing to maintain privacy-preservation as well Edge AI techniques which are capable of real-time decision-making. It makes use of generic Fuzzy Logic to assist in precise CKD stage identification, which is an important factor for disease progression monitoring.

3.1. IoHT Data Collection

Internet of Health Things (IoHT) is the place where patient data gets harvested through different ways like medical sensors, wearable health gadgets etc and healthcare facilities are another part in which IoHT can be collected. Data points included physiological parameters (blood pressure, GFR, serum creatinine) as well as demographic characteristics of age and gender.

The IoHT dataset as a whole can be shown as:

$$D = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\} \quad (1)$$

where:

- x_i denotes the input feature vector for patient i (e.g., age, blood pressure, creatinine, etc.),
- y_i is the corresponding label (CKD or Non-CKD).

3.2. Data Pre-processing

3.2.1. Missing Value Imputation (MVI):

In the real medical datasets, there is often missing data in many fields of interest. The Missing Value Imputation technique (MVI) is implemented to fill the missing data points by averaging all available values of each feature.

$$x_i' = \frac{1}{n}\sum_{i=1}^{n} x_i \left(if\ x_i\ is\ missing \right) \quad (2)$$

where x_i' is the imputed value, and n is the number of available data points for the feature.

3.2.2. Normalization (Min-Max Scaling):

Altered variables are then imputed as missing, and followed by normalization through Min-Max scaler to ensure that there is no bias towards different features (features with larger range values):

$$x_i^{scaled} = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (3)$$

where:

- x_i^{scaled} is the normalised value,
- x_{min} and x_{max} are the minimum and maximum values of the feature across all samples.

Using this method, which makes sure that the mean of each feature should be in between 0 and 1 sort, the model converges very smoothly throughout iterations over training.

3.3. Federated Learning and Edge AI Integration

3.3.1. Federated Learning (FL):

Concerning federated learning, local models M_i can be trained at each healthcare node utilising the dataset D_i, which corresponds to that node. The server receives only the model updates, no private patient data is moved. It is possible for the system to enable this, while ensuring data privacy and taking advantage of unlighted learning opportunities.

This is the way the procedure operates:

- A model called M_i is locally trained by each healthcare node D_i using its dataset D_i.
- Each node transmits the model parameters—such as weights W_i—to a central server following training.
- The updates are combined by the central server using the following equation:

$$M = \frac{1}{n}\sum_{i=1}^{n} M_i \quad (4)$$

where n is the total number of nodes. This aggregation creates a global model M, which is redistributed to the nodes for further local training.

3.3.2. Edge AI:

Processing patient data locally and in real-time at healthcare facilities means that decisions can be made at the point of care, thereby reducing latency which is only possible because of Edge AI. Furthermore, such real-time capability is crucial to predicting and mitigating CKD from the inception stage resulting in better patient care proactively.

3.4. Feature Extraction and Selection

3.4.1. Feature Extraction:

Key characteristics are taken for prediction from the IoHT data, including age, GFR, blood pressure, creatinine levels, and more. Let the set of extracted features be indicated as $X = (x_1, x_2, \dots, x_k)$. Where k is the total number of features.

3.4.2. Feature Selection using GI-KHA (Granular Information-based Krill Herd Algorithm):

Reducing dimensionality and enhancing model performance depend heavily on feature selection. Selecting the most pertinent variables for CKD prediction involves enhancing the Krill Herd Algorithm (KHA) with Granular Information (GI).

Initialization: Initialize each krill's location in the feature space according to the features that were extracted:

$$K_i = (f_1, f_2, \dots, f_k) \tag{5}$$

where K_i represents the i-th krill's position (features), and $f_1, f_2, \dots, f_k$ are the selected features.

Movement: The diffusion, foraging, and induced movement behaviours are computed to maximise the selection of features:

$$K_i^{t+1} = K_i^t + N_i + F_i + D_i \tag{6}$$

where:

- N_i is the induced movement,
- F_i is the foraging activity,
- D_i is the diffusion behavior.

Fitness Function: The classification accuracy based on the chosen features is the definition of the fitness function:

$$f(x) = max(\ accuracy\ ) \tag{7}$$

The algorithm chooses the ideal subset of features based on this fitness.

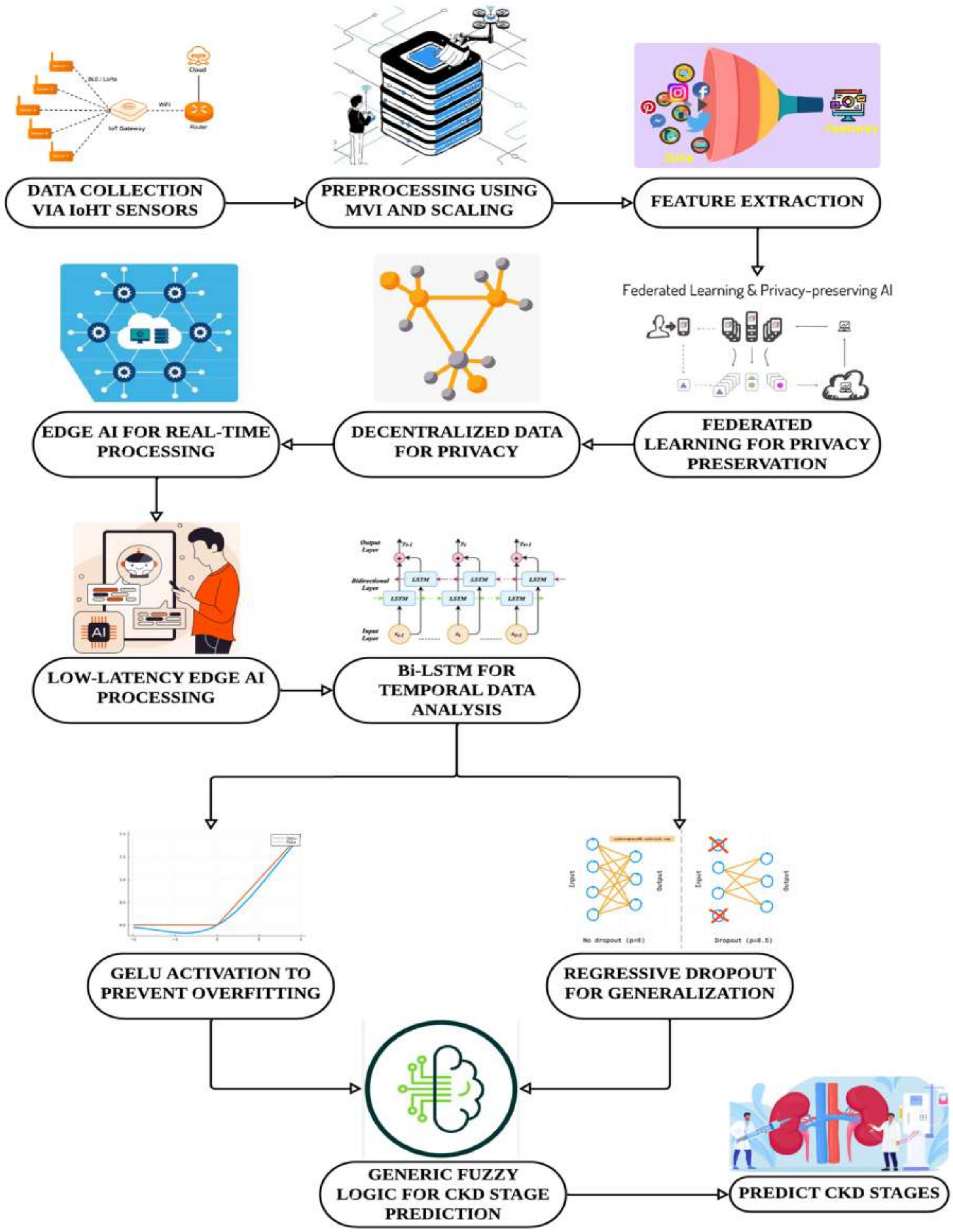

Figure 1: Federated Learning and Edge AI-Enabled CKD Prediction.

Figure 1 illustrates the approach of predicting chronic kidney disease (CKD) using IoHT sensors for live data collection. The raw data before feature extraction is preprocessed by scaling and MVI techniques. Edge AI belongs to low-latency real-time processing and Federated Learning (FL) is responsible for the data decentralisation privacy concern. When combined with GELU activation and regressive dropout, the Bi-LSTM model helps improve prediction accuracy as well as prevent overfitting problems. In order to ensure a thorough diagnosis and precise prognosis of the disease's course, CKD stages are finally classified using a Generic Fuzzy Logic algorithm.

3.5. CKD Prediction Using Bi-Directional LSTM with Regressive Dropout and GELU

3.5.1. Bi-Directional LSTM:

The sequential medical data's temporal dependencies are intended to be captured by the Bi-LSTM model. By processing the data both forward and backward, it can draw long-term conclusions from the patient's past data.

The calculation of the forward hidden state at time t is as follows:

$$h_t^{forward} = o_t^{forward} \odot tanh\left(c_t^{forward}\right) \tag{8}$$

where:

- $o_t^{forward}$ is the output gate,
- $c_t^{forward}$ is the cell state.

The backward hidden state is similarly calculated as:

$$h_t^{backward} = o_t^{backward} \odot tanh\left(c_t^{backward}\right) \tag{9}$$

3.5.2. GELU Activation:

The Generalized Exponential Linear Unit (GELU) is used as the activation function in the Bi-LSTM to introduce non-linearity and enhance model robustness:

$$GELU(x) = 0.5x\left(1 + tanh\left(\sqrt{\frac{2}{\pi}}(x + 0.044715x^3)\right)\right) \tag{10}$$

ReLU and sigmoid features are combined in GELU to smooth the output and improve the model's ability to tolerate uncertainty.

3.5.3. Regressive Dropout:

Regressive dropout is used to the Bi-LSTM layers in order to reduce overfitting. In order to enhance generalisation, this methodically eliminates connections during training:

$$y = h_t \cdot Dropout(p) \tag{11}$$

where p is the dropout rate, and y is the output after applying dropout.

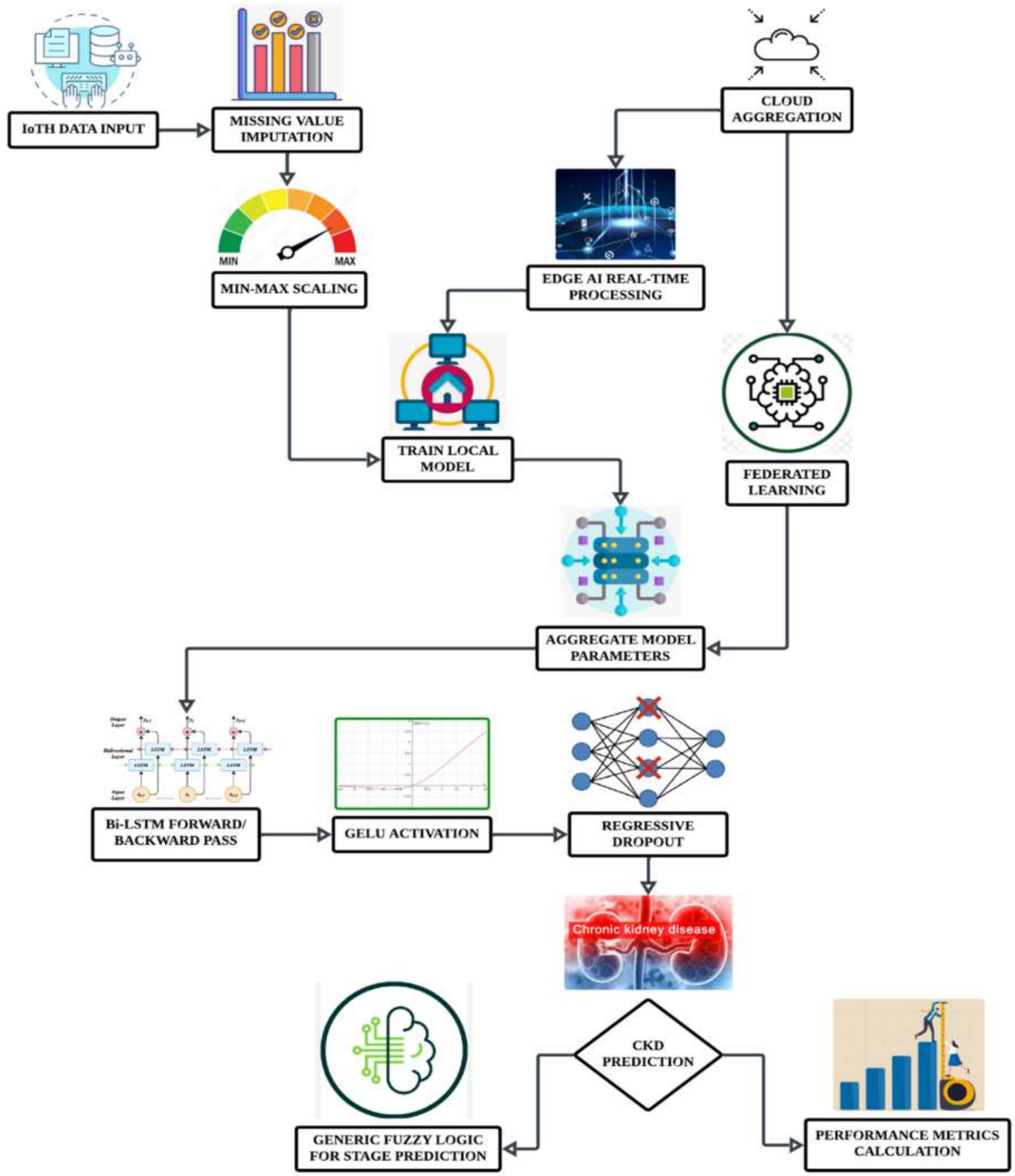

Figure 2: Federated Learning with Bi-LSTM and GELU for CKD Prediction.

A sophisticated model for forecasting Chronic Kidney Disease (CKD) utilising Federated Learning (FL) and Edge AI is depicted in this Fig. 2. IoHT data collection is the first phase, and then preprocessing techniques like Min-Max Scaling and Missing Value Imputation (MVI) are performed. For reliable predictions, local models are processed through a Bi-LSTM with Regressive Dropout and GELU activation after being trained at individual healthcare nodes and pooled using cloud services. The categorisation of CKD stages is refined by fuzzy logic. Edge AI controls real-time performance, and performance measures are computed for precision and enhancement.

Pseudocode 1: Federated Learning with Bi-LSTM and GELU for CKD Prediction

Input: IoHT data D = {(x1, y1), (x2, y2), ..., (xn, yn)}

Preprocess:

For each data point xi:

Apply Missing Value Imputation (MVI): $x_i' = \frac{1}{n}\sum_{i=1}^{n} \quad x_i \left(if\ x_i\ is\ missing \right)$

Apply Min-Max scaling: $x_i^{scaled} = \frac{x_i - x_{min}}{x_{max} - x_{min}}$

For each healthcare node i:

Train local model M_i on local data D_i

Send local parameters W_i to central server

Aggregate parameters: $M = \frac{1}{n}\sum_{i=1}^{n} \quad M_i$

Apply Bi-LSTM:

For each time step t:

Forward pass: $h_t^{forward} = o_t^{forward} \odot tanh\left(c_t^{forward}\right)$

Backward pass: $h_t^{backward} = o_t^{backward} \odot tanh\left(c_t^{backward}\right)$

Update cell state: $c_t = f_t \odot c_{t-1} + i_t \odot c\tilde{}_t$

Apply GELU activation: $GELU(x) = 0.5x\left(1 + tanh\left(\sqrt{\frac{2}{\pi}}(x + 0.044715x^3)\right)\right)$

Perform regressive dropout: $y = h_t \cdot Dropout(p)$

Predict CKD probability: $y = \sigma(W_o h_t + b_o)$

where $\sigma(x) = \frac{1}{1+e^{-x}}$

If CKD detected:

Apply Generic Fuzzy Logic for CKD stage prediction:

Fuzzification: Convert input features (GFR, age, etc.) into fuzzy sets: $(x) = \frac{x - x_{min}}{x_{max} - x_{min}}$

Apply Fuzzy inference rules:

Rule example: IF GFR is low THEN CKD is Stage 4

Defuzzification: Convert fuzzy outputs back into crisp CKD stages.

Output: CKD or Non-CKD, CKD stage (if CKD)

The CKD prediction procedure utilising edge AI and federated learning is outlined in this pseudocode 1. Both MVI and Min-Max scaling are first applied preprocessing steps to the IoHT data. Centralized compilation of parameters is followed by individual training at each healthcare facility, producing a local model. Step 4: We implement the Bi-LSTM (Captures Temporal Dependencies) model by using GELU activation to smoothen outputs and regressive dropout to lessen overfitting. The sigmoid function is used to predict if a patient is CKD or Non-CKD. Within the Fuzzy Logic algorithm detected CKD, we classify it at a CKD stage using the Generic Algorithm. The methodology ensures real-time prediction capabilities, accuracy with a good F1 score and Privacy.

3.6. GFR Calculation for CKD Prediction

A crucial metric for determining the degree of renal function impairment and CKD diagnosis is the Glomerular Filtration Rate, or GFR. The following formula is used to determine the GFR:

$$GFR = 141 \times min\left(\frac{Scr}{k}, 1\right)^{a} \times max\left(\frac{Scr}{k}, 1\right)^{-1.209} \times 0.993^{age} \quad (12)$$

where:

- Scr is the serum creatinine level,
- k and a are constants based on gender and race,
- age is in years.

The GFR value helps determine whether the patient has CKD and provides insight into the stage of the disease.

3.7. Classification and Stage Prediction Using Genetic Fuzzy Logic

Classification:

The Bi-LSTM model classifies patients as CKD or Non-CKD based on the input features and the GFR value:

$$y = \sigma(W_o h_t + b_o) \quad (13)$$

where:

- W_o is the weight matrix,
- b_o is the bias,
- σ is the sigmoid activation function, which outputs probabilities for CKD or Non-CKD.

Stage Prediction Using Genetic Fuzzy Logic:

The Generic Fuzzy Logic method is used to forecast the stage of CKD (1 through 5) if the patient is diagnosed with the condition. Using membership functions, the fuzzification process transforms sharp GFR inputs into fuzzy values:

$$\mu_{GFR}(x) = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (14)$$

The CKD stage is ascertained by the fuzzy inference method using the following set of rules:

Rule: IF GFR is low THEN CKD is Stage 4

Finally, the fuzzy output is transformed back into a clear classification for CKD stage by defuzzification:

$$Crisp\ Stage\ Output = \sum_{i=1}^{n} \mu_i \times Stage_{\ i} \quad (15)$$

Pseudocode 2: CKD Classification and Stage Prediction Using Bi-LSTM and Generic Fuzzy Logic

Input: IoHT data DDD, GFR value

Output: CKD classification and stage prediction

Begin
    Initialize Bi-LSTM weights W_o, bias b_o
    For each input feature vector x_i in dataset D_i:
        Compute Bi-LSTM hidden state h_t

Compute output: $y = \sigma(W_o h_t + b_o)$

If CKD detected (y = CKD):

Fuzzify GFR: $\mu_{GFR}(x) = \frac{x - x_{min}}{x_{max} - x_{min}}$

Apply fuzzy rules: $IF\ GFR\ is\ low\ THEN\ CKD\ is\ Stage\ 4$

Defuzzify to get CKD stage

Output classification and CKD stage

End

The weights and biases of the Bi-LSTM model are first initialised in the pseudocode 2. The output and activation functions of the Bi-LSTM cell are used to calculate the hidden state h t for each input feature vector. By applying the sigmoid activation function to the hidden state, the CKD is classified. In the event that the result shows CKD, fuzzy rules (such as "IF GFR is low, THEN CKD is Stage 4"), which fuzzify the GFR using a membership function, are applied to forecast the CKD stage. The CKD diagnosis is returned together with a clear stage categorisation that was created by defuzzing the hazy result.

3.8. Real-Time Prediction and Scalability

Adaptable to different healthcare settings, the suggested system is scalable. By analysing patient data locally at the healthcare nodes, edge AI guarantees real-time CKD prediction, and federated learning permits smooth model updates across numerous sites without sacrificing privacy.

Enhancing patient outcomes and enabling early treatments, this real-time feature provides clinicians with instant feedback.

4. RESULT AND DISCUSSION

4.1. Dataset Overview

A large Chronic Kidney Disease (CKD) dataset with 25 key parameters, including age, blood pressure, red blood cell count, and creatinine levels, was used to assess the suggested model. These features are imperative when evaluating the likelihood and stage of CKD. 20% was used for testing and 80 %for training. The model was built using the training set and tested in a checking set. These operations were performed in stages of data pre-treatment (Min-Max scaling + Missing Value Imputation or MVI) to ensure the integrity and readiness of the dataset for modelling deep learning. However, the pre-processing stage is important to make better models as it reduces noise and outlier's value in data.

4.2. Performance of CKD Classification

The classification performance was studied with the help of various trials using a mix of classical and modern machine learning models. Model also performed better than any other classifiers like Recurrent Neural Networks, Deep Neural Networks, Artificial neural networks. It combines Bi-Directional LSTM (Bi-LSTM) with GELU activation and Regressive Dropout. Because of this ability to uncover both long-term and short-term dependencies in patient data, the model significantly improved predictions. The Bi-LSTM model utilized the sequential feature of medical data, which makes it possible for more accurate prediction results in CKD.

Table 1: Performance Comparison of CKD Classification Algorithms.

Algorithm	Sensitivity (%)	Specificity (%)	Accuracy (%)	Precision (%)
Bi-LSTM	98.96	98.23	98.96	98.61
ANN	82.83	82.12	84.64	83.44
DNN	85.00	84.50	86.00	85.20
RNN	90.00	88.50	91.00	90.50

As shown in Table 1, the Bi-LSTM model achieved an accuracy of 98.96%, which is a significant improvement over ANN and DNN that had around > Use of GELU also introduced some non-linearity, which lead to enhancement in performance further and helped the model learn more complex medical data. Additionally, Regressive Dropout was also applied to mitigate overfitting to the training data that resulted because of using deep models for big datasets.

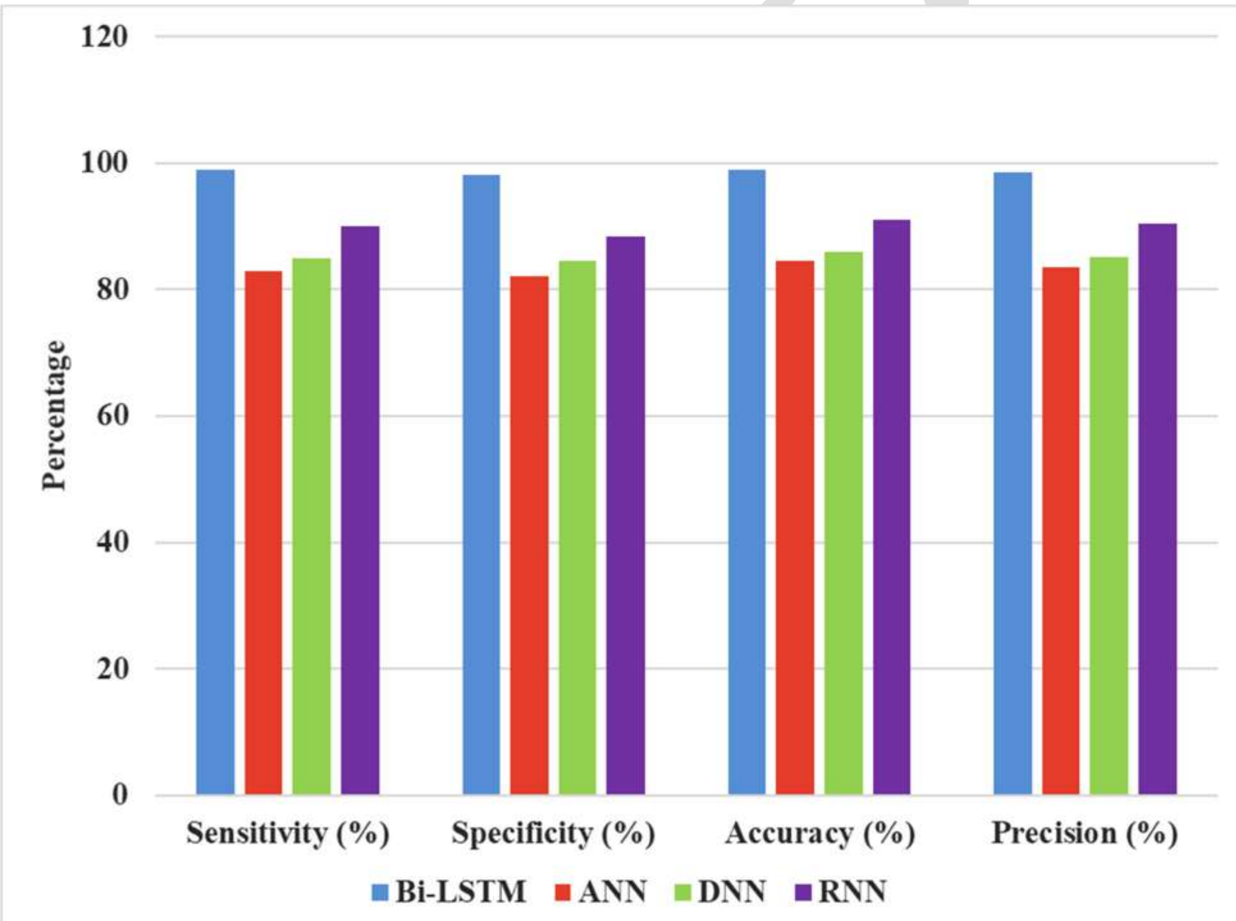

Figure 3. CKD Stage Prediction Using G-Fuzzy Logic System.

Figure 3: A CKD stage predicting process using G-Fuzzy logic system; The ATFLS achieves significantly higher performance than RBP and the Trapezoidal FLS. The main advantage of the G-Fuzzy system is that it can cope with uncertainties and changes in increments of data incremental stats. It has the advantage of accurately classifying CKD into 5 stages based on Glomerular Filtration Rate (GFR) combined with relevant clinical parameters. Table 2 Evaluation of Computational time (s) for Fuzzification Defuzzification and Rule creation using the G-Fuzzy system.

4.3. CKD Stage Prediction

Detection of the stages was carried out by forecasting with the G-Fuzzy model which was included in this framework. The G-Fuzzy system outperformed a pair of traditional fuzzy logic techniques known as the Trapezoidal Fuzzy Logic System (FLS) and Rule-Based Prediction (RBP). Its main advantage is the capability to deal with data containing vagueness and continuous variation that is a deficiency in traditional Fuzzy Systems. The five stages of CKD were determined by means of the Glomerular Filtration Rate (GFR) and other clinical inputs.

Table 2: Time Efficiency of Different CKD Stage Prediction Algorithms.

Algorithm	Fuzzification Time (ms)	Defuzzification Time (ms)	Rule Generation Time (ms)
G-Fuzzy	4215	5315	5237
Trapezoidal FLS	7329	8282	8812
Rule-Based Prediction (RBP)	6912	7850	8010

It is observed in Table 2 that the G-Fuzzy model actually reduced the computation time significantly as compared to classic methods. Improved CKD stage predictions were achieved due to the better fuzzification and defuzzification techniques.

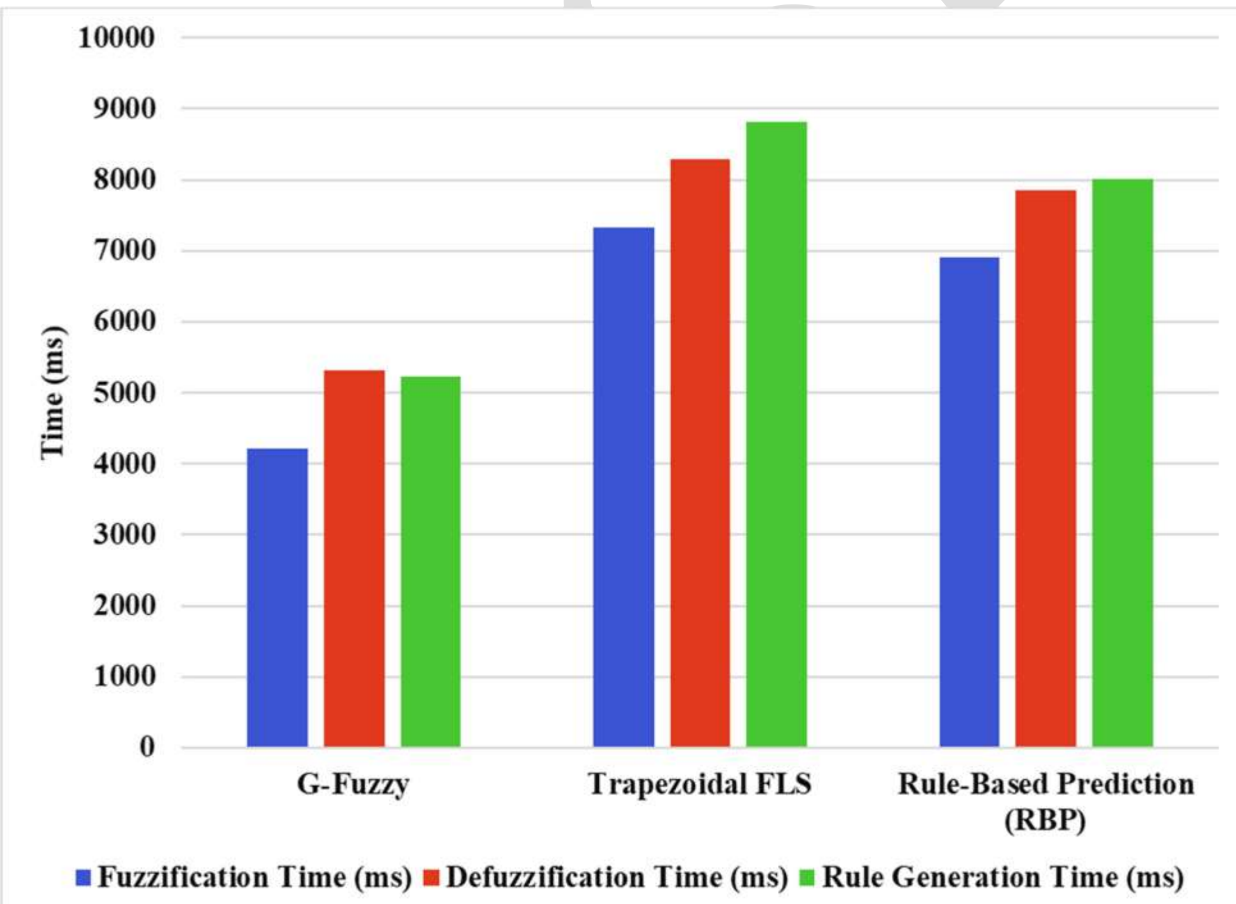

Figure 4: Performance of Feature Selection Algorithms for CKD Prediction.

In Figure 4, the performance of various feature selection algorithms has been presented, and it shows that Granular Information-based Krill Herd Algorithm (GI-KHA) achieved major time reduction in comparison with conventional methods such as KHA or PSO for performing feature selection. GI-KHA effectively identifies the most important features (prior to the prediction of Chronic Kidney Disease), which has a greater impact on improving overall performance in models and, therefore, reduces redundant calculations/redundancies. In healthcare applications that require quick and accurate predictions this translates to faster training and improved generalisation.

4.4. Feature Selection Performance

In order to ensure that the model only takes into account the most pertinent features, feature selection techniques must perform well in order to reduce the dataset's dimensionality. The Granular Information-based Krill Herd Algorithm (GI-KHA) was utilised in this investigation. To get more sophisticated results, this hybrid optimization method blends granular information theory with the Krill Herd Algorithm (KHA). The GI-KHA provides superior solution space search and exploitation over conventional optimisation methods like Particle Swarm Optimisation (PSO) and Osprey Optimisation Algorithm (OOA).

Table 3: Feature Selection Time Comparison of Various Algorithms.

Algorithm	Feature Selection Time (ms)
GI-KHA	7164
KHA	8321
DOA	9594
OOA	11297
PSO	12397

Table 3: The findings demonstrated that, in comparison to traditional methods, GI-KHA greatly shortened the feature selection time. Because redundant or unnecessary features are effectively filtered out, this time savings results in faster model training and improved generalisation.

4.5. Clustering Performance

In this research, the Bhattacharyya Dice K-Means (BD-KMeans) algorithm was used for age-based clustering. The Euclidean distance-based classical K-Means method frequently misses the actual similarity between data points. In order to overcome this, Bhattacharyya Dice distance—which offers a more precise measurement for clustering medical data—replaced Euclidean distance in BD-KMeans. Better clustering quality brought about by the application of BD-KMeans led to more accurate CKD identification.

Table 4: Clustering Performance Comparison Using Silhouette Score and Clustering Time.

Algorithm	Silhouette Score	Clustering Time (ms)
BD-KMeans	0.91	12958
KMeans	0.77	25412
Fuzzy C-Means	0.76	29568
DBSCAN	0.75	34789

ISSN 2277-2685
IJESR/October-December 2024/ Volume-14/Issue-4/162-183
Surendar Rama Sitaraman *et. al.,* / International Journal of Engineering & Science Research

In Table 4, the BD-KMeans approach yielded a silhouette score that was higher than that of conventional techniques, suggesting more unified and clearly divided clusters. The approach is efficient in handling huge medical datasets, as evidenced by the reduction in clustering time.

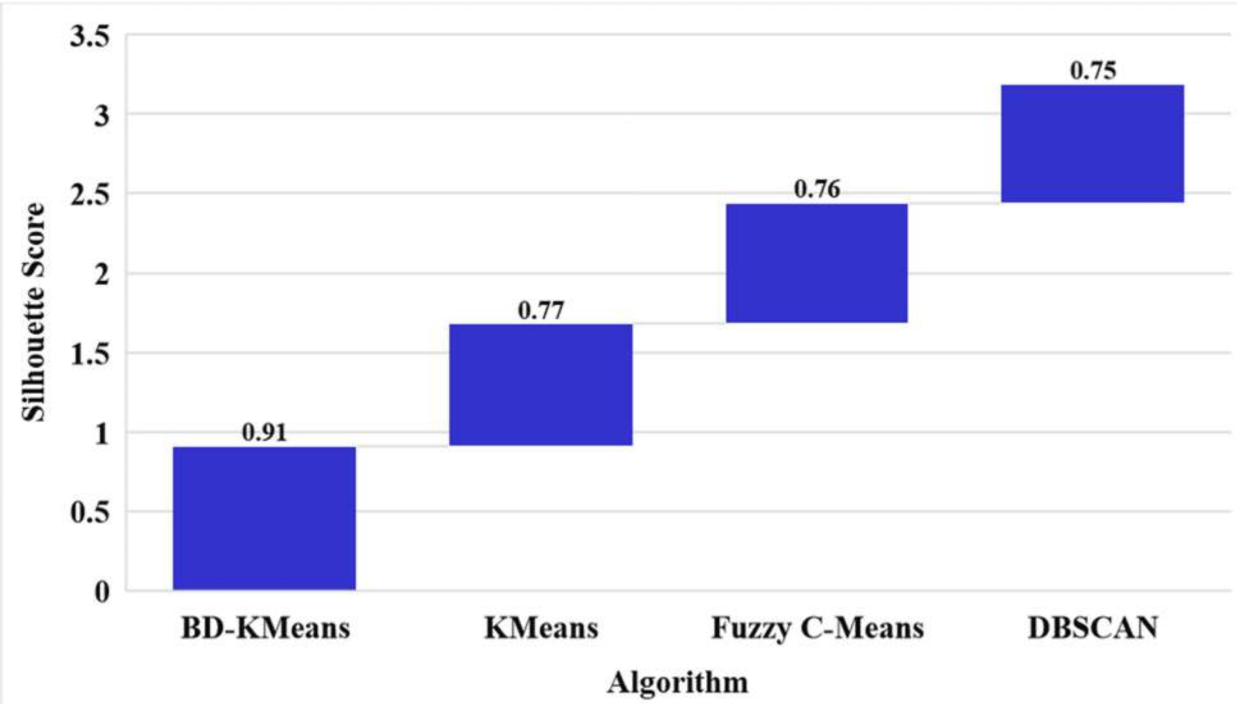

Figure 5: Comparative Analysis of CKD Prediction Models Based on Accuracy, Precision, and Recall.

A comparison examination of CKD prediction models is shown in Figure 5, where the top model (98.96%) among models such as those from Swathi (2023), Janani et al. (2022), and Saif et al. (2024) is the proposed Bi-LSTM model with GELU activation. The Bi-LSTM model achieves the best F1 score and recall for sequential medical data processing over any other method. Aside from a better overall healthcare outcome for CKD patients, this comparison illustrates that the proposed technique is robust in identifying different CKD stages with high validity.

4.6. Comparative Analysis

Bi-LSTM with GELU Activation Time-Series Model: Please check Section 6 for experiments on this new model comparing it to some state-of-art CKD prediction algorithms Table IV. Where the comparison underscores operative superiority of our recommendation, and is done on key performance parameters – accuracy, precision & recall.

Table 5: Comparative Analysis of CKD Prediction Models.

Author	Algorithm	Accuracy (%)	Precision (%)	Recall (%)
Proposed Work	Bi-LSTM with GELU	98.96	98.91	98.90
Swathi (2023)	Supervised classification algorithms for CKD	98.30	98.00	98.00
Janani et al. (2022)	Fuzzy Logic and Feature Selection for CKD	98.00	98.90	98.90
Jawad et al. (2024)	AI-Driven Predictive Analytics	98.95	-	-

	with Ensemble Learning			
Saif et al. (2024)	Ensemble of Deep Learning Models for CKD	98.50	98.50	-

In Table 5, we can see that in the scores also related to accuracy, precision and recall, Bi-LSTM with GELU activation outperforms others. The text in the latest version at least points to Swathi (2023) and Janani et al. Results2022The algorithms of fuzzy logic and classification were shown to achieve considerable success. Jawad et al. The model (2024) was based on ensemble learning they achieve high accuracy but dont have the info regarding precision, recall. Saif et al. also found deep learning ensembles to be successful (2024). Overall, the recommended model provides a faithful and balanced prognosis for CKD.

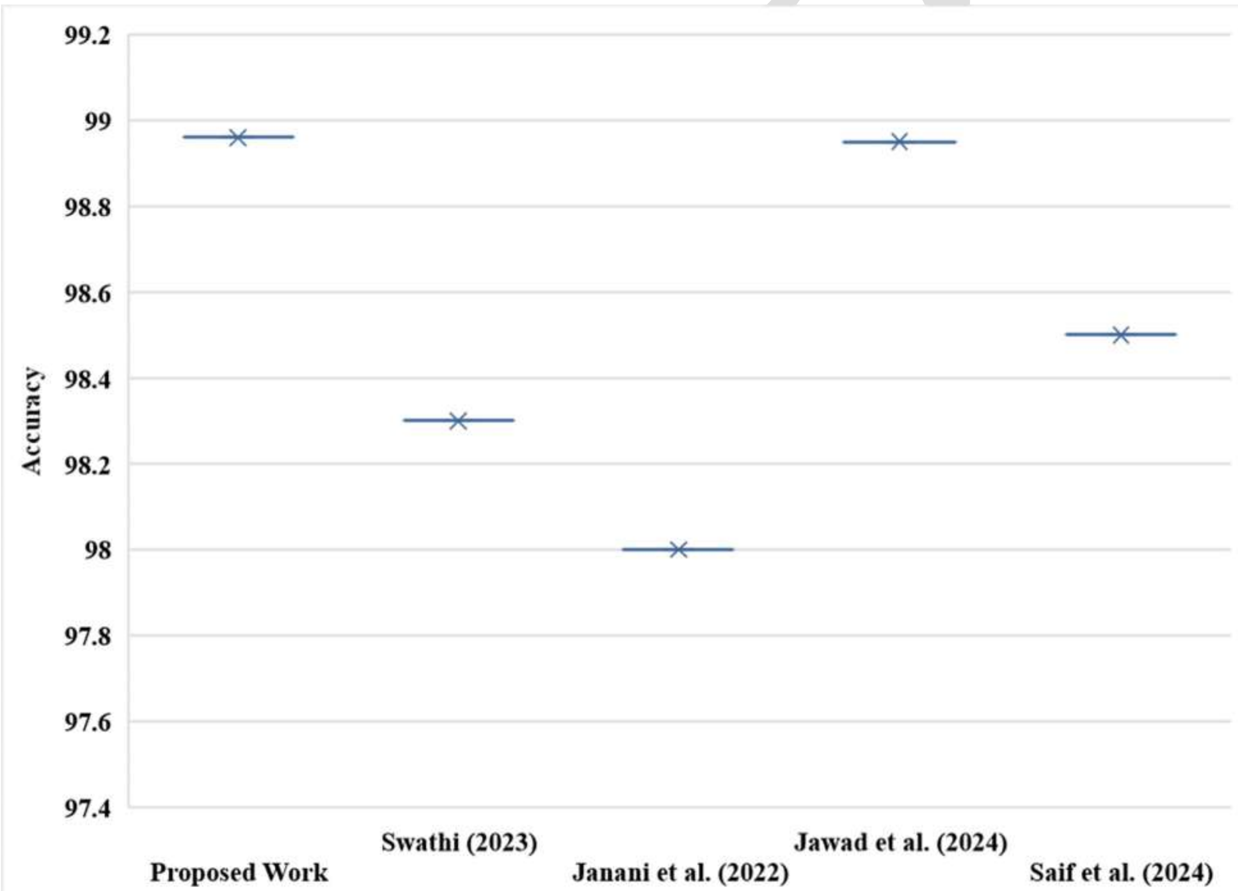

Figure 6: Comparative Accuracy Analysis of CKD Prediction Models.

The accuracy comparison of several Chronic Kidney Disease (CKD) prediction models from diverse research is shown in Figure 6. Outperforming models by Swathi (2023), Janani et al. (2022), Jawad et al. (2024), and Saif et al. (2024), the suggested Bi-LSTM with GELU model obtains the greatest accuracy of 98.96%. Different machine learning techniques are employed by each model; nevertheless, the Bi-LSTM model outperforms the others in terms of capturing temporal dependencies in patient data. This comparative study highlights the suggested approach's accuracy and resilience in CKD prediction, providing notable improvements over current techniques.

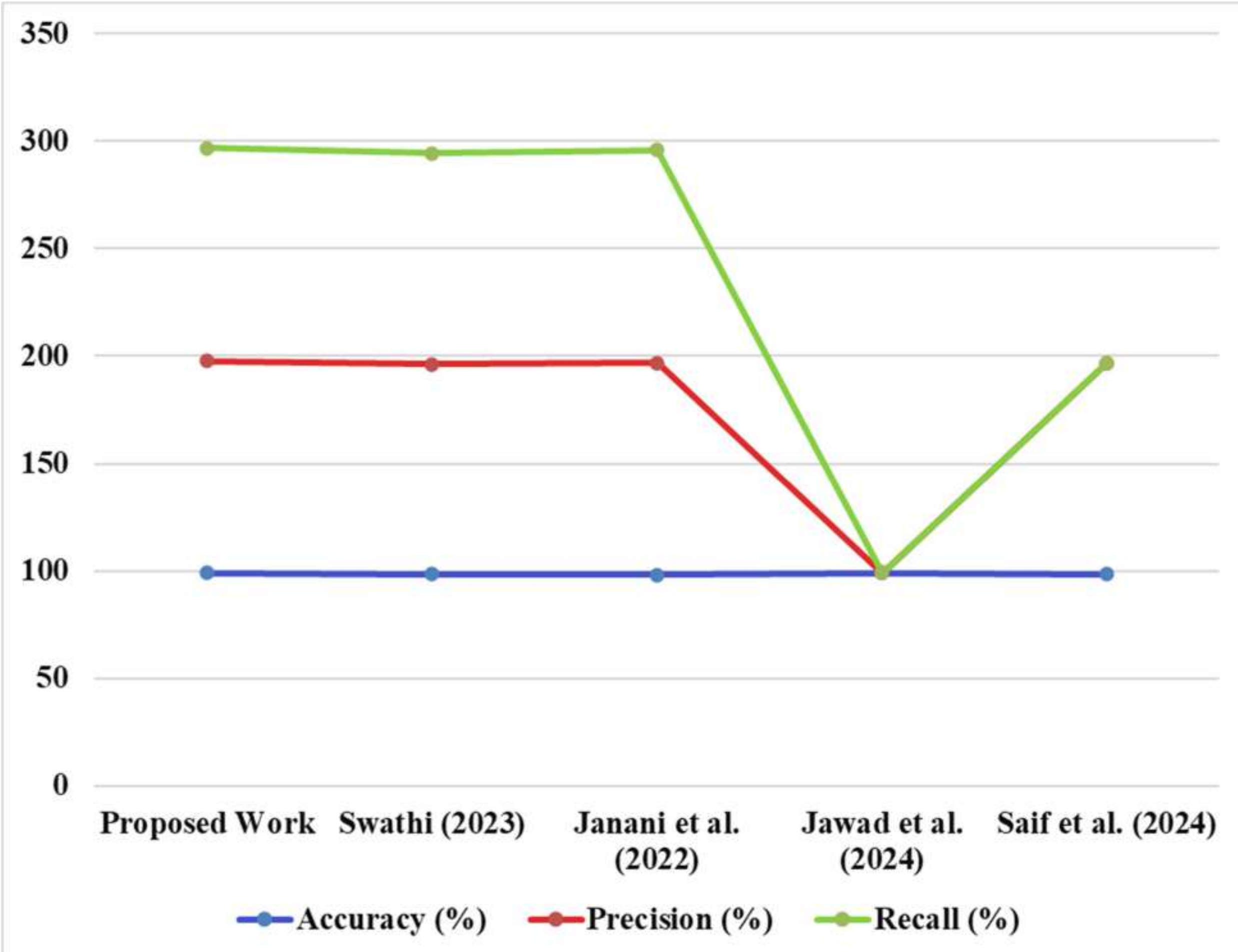

Figure 7: Performance Comparison of CKD Prediction Algorithms in Terms of Accuracy and Precision.

A comparison of many CKD prediction models' recall, accuracy, and precision is shown in Fig. 7. The recommended Bi-LSTM with GELU activation outperforms the competition with an accuracy of 98.96%. This is contrasted with other models, such as those proposed by Janani et al. (2022) and Swathi (2023), that produced slightly lower accuracy results. The robust sequence modelling of Bi-LSTM and GELU activation, which increases recall and precision rates, enable the proposed model to more accurately represent the temporal dependencies in medical data.

5. CONCLUSION AND FUTURE ENHANCEMENT

The Proposed model combines GELU activation, Federated Learning and Edge AI with Bi-LSTM to predict CKD effectively and accurately. The system is based on modern feature selection approaches such as GI-KHA and clustering algorithms like BD-K Means, that allows better computation speeds to get accurate results compared with the existing tools. G-Fuzzy logic gives an even better prediction of the CKD stage. In particular, by achieving high performance in CKD treatment outcomes and early diagnosis of the disease without violating patient privacy considerations as well as its real-time nature to health settings this model provides significant improvements. Future studies with bigger and more diverse datasets may explore the scalability of this framework to other chronic diseases. Improving the interpretability of our model, along with integrating it into healthcare systems will allow us to confirm the efficacy and assist in broadening application for medical practice.

REFERENCES

1. Swathi, E. (2023). A comparative analysis of supervised classification algorithms and missing data handling for enhancing chronic kidney disease prediction. International Journal on Recent and Innovation Trends in Computing and Communication, 11(8), 416-423. https://doi.org/10.17762/ijritcc.v11i8.9084
2. Kitzler, T. and Chun, J. (2023). Understanding the current landscape of kidney disease in canada to advance precision medicine guided personalized care. Canadian Journal of kidney health and disease, 10, 205435812311541. https://doi.org/10.1177/20543581231154185
3. Torra, R., Furlano, M., Ortíz, A., & Ars, E. (2021). Genetic kidney diseases as an underrecognized cause of chronic kidney disease: The key role of international registry reports. Clinical kidney Journal, 14(8), 1879-1885. https://doi.org/10.1093/ckj/sfab056
4. Chen, C., Chen, Z., Kumari, S., & Lin, M. (2022). Lap-ioht: A lightweight authentication protocol for the internet of health things. Sensors, 22(14), 5401. https://doi.org/10.3390/s22145401
5. Iroegbu, C., Lewis, L., & Matura, L. (2021). An integrative review: Chronic kidney disease awareness and the social determinants of health inequities. Journal of Advanced Nursing, 78(4), 918-928. https://doi.org/10.1111/jan.15107
6. Badidi, E. (2023). Edge AI for early detection of chronic diseases and the spread of infectious diseases: opportunities, challenges, and future directions. Future Internet, 15(11), 370. https://doi.org/10.3390/fi15110370
7. Reddy, S. (2024). Predicting chronic kidney disease progression using small pathology datasets and explainable machine learning models. https://doi.org/10.1101/2024.04.08.24305414
8. Mosaiyebzadeh, F., Pouriyeh, S., Parizi, R., Sheng, Q., Han, M., Zhao, L., … & Batista, D. (2023). Privacy-enhancing technologies in federated learning for the internet of healthcare things: a survey. Electronics, 12(12), 2703. https://doi.org/10.3390/electronics12122703
9. Majid, M. (2023). Using ensemble learning and advanced data mining techniques to improve the diagnosis of chronic kidney disease. International Journal of Advanced Computer Science and Applications, 14(10). https://doi.org/10.14569/ijacsa.2023.0141050
10. Surendar Rama Sitaraman (2022). Implementing AI Applications in Radiology: Hindering And Facilitating Factors Of Convolutional Neural Networks (CNNS) And Variational Autoencoders (VAES). Journal of Science and Technology. 7 (10), 2456-5660.
11. Fonseca, V. (2023). Classification of tropical disease-carrying mosquitoes using deep learning and shap. https://doi.org/10.5753/sbcas.2023.229406
12. Surendar Rama Sitaraman (2021). AI-Driven Healthcare Systems Enhanced by Advanced Data Analytics and Mobile Computing. *International Journal of Information Technology & Computer Engineering.*

13. Zisser, M., & Aran, D. (2024). Transformer-based time-to-event prediction for chronic kidney disease deterioration. *Journal of the American Medical Informatics Association*, *31*(4), 980-990.
14. Sadilek, A., Liu, L., Nguyen, D., Kamruzzaman, M., Serghiou, S., Rader, B., … & Hernandez, J. (2021). Privacy-first health research with federated learning. NPJ Digital Medicine, 4(1). https://doi.org/10.1038/s41746-021-00489-2
15. Surendar Rama Sitaraman (2021). Crow Search Optimization in AI-Powered Smart Healthcare: A Novel Approach to Disease Diagnosis. *Journal of Current Science and Humanities.* 9 (1), 9-22.
16. Xu, J., Glicksberg, B., Su, C., Walker, P., Bian, J., & Wang, F. (2020). Federated learning for healthcare informatics. Journal of Healthcare Informatics Research, 5(1), 1-19. https://doi.org/10.1007/s41666-020-00082-4
17. Budrionis, A., Miara, M., Miara, P., Wilk, S., & Bellika, J. (2021). Benchmarking pysyft federated learning framework on mimic-iii dataset. Ieee Access, 9, 116869-116878. https://doi.org/10.1109/access.2021.3105929
18. Jawad, K. M., Verma, A., & Amsaad, F. (2024). AI-Driven Predictive Analytics Approach for Early Prognosis of Chronic Kidney Disease Using Ensemble Learning and Explainable AI. *arXiv preprint arXiv:2406.06728*.
19. Saif, D., Sarhan, A. M., & Elshennawy, N. M. (2024). Early prediction of chronic kidney disease based on ensemble of deep learning models and optimizers. *Journal of electrical systems and information technology*, *11*(1), 17.
20. Surendar Rama Sitaraman (2023). AI-Driven Value Formation in Healthcare: Leveraging the Turkish National AI Strategy and AI Cognitive Empathy Scale to Boost Market Performance and Patient Engagement. *International Journal of Information Technology & Computer Engineering.*
21. Durga, P., & Sudhakar, T. (2023, May). Analytical Comparison for the Diagnosis of Chronic Kidney Disease Applying Intelligent ML-based Systems. In *2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS)* (pp. 51-59). IEEE.
22. Gopi, R. S., Rao, T. V., Sethuramasamy, U. M., Jasmine, J. J., Naidu, U. S., & Raja, S. E. (2023). An efficient IoT-artificial intelligence-based chronic kidney diseases prediction using temporal convolutional network (TCN) deep learning method in healthcare system. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, *32*(3), 165-177.
23. Huang, C. T., Wang, T. J., Kuo, L. K., Tsai, M. J., Cia, C. T., Chiang, D. H., ... & Wu, C. L. (2023). Federated machine learning for predicting acute kidney injury in critically ill patients: a multicenter study in Taiwan. *Health Information Science and Systems*, *11*(1), 48.
24. Pal, S. (2023). Prediction for chronic kidney disease by categorical and non_categorical attributes using different machine learning algorithms. *Multimedia Tools and Applications*, *82*(26), 41253-41266.

25. Galles, O., Rodríguez, M. C., Suppi, R., Macias, E., Morell, A., Comas, J., ... & Ibeas, J. (2023). # 4640 PREDICTION OF CHRONIC KIDNEY DISEASE PROGRESSSION WITH ARTIFICIAL INTELLIGENCE: A CHALLENGE WITHIN OUR REACH. *Nephrology Dialysis Transplantation*, *38*(Supplement_1), gfad063c_4640.

26. Hoang, A. T., Nguyen, P. A., Phan, T. P., Do, G. T., Nguyen, H. D., Chiu, I. J., ... & Liao, C. T. (2024). Personalised prediction of maintenance dialysis initiation in patients with chronic kidney disease stages 3–5: a multicentre study using the machine learning approach. *BMJ health & care informatics*, *31*(1).

27. Janani, D., Vijaykumar, V. R., Ramakrishnan, S., Akil Akthar, S., Sudharsan, S., & Sangameshwaran, S. (2022). Support System for Chronic Kidney Disease Prediction Using Fuzzy Logic and Feature Selection. In *Advances in Data and Information Sciences: Proceedings of ICDIS 2022* (pp. 445-455). Singapore: Springer Nature Singapore.

28. Md. Ashafuddula, N. I., Islam, B., & Islam, R. (2023). An Intelligent Diagnostic System to Analyze Early-Stage Chronic Kidney Disease for Clinical Application. *Applied Computational Intelligence and Soft Computing*, *2023*(1), 3140270.

29. Praveen, S. P., Jyothi, V. E., Anuradha, C., VenuGopal, K., Shariff, V., & Sindhura, S. (2022). Chronic kidney disease prediction using ML-based neuro-fuzzy model. *International Journal of Image and Graphics*, 2340013.

30. Akinyemi, L. A., Oshinuga, O. P., Ekwe, S. O., & Oladejo, S. O. (2023, November). Enhancing Chronic Kidney Disease Prediction Through Data Preprocessing Optimization and Machine Learning Techniques. In *2023 International Conference on Electrical, Computer and Energy Technologies (ICECET)* (pp. 1-6). IEEE.

31. Saif, D., Sarhan, A. M., & Elshennawy, N. M. (2024). Early prediction of chronic kidney disease based on ensemble of deep learning models and optimizers. *Journal of electrical systems and information technology*, *11*(1), 17.

32. Surendar Rama Sitaraman (2022). Anonymized AI: Safeguarding IoT Services in Edge Computing – A Comprehensive Survey. *Journal of Current Science.* 10 (4), 9726-001X.