Article

A Deep Neural Network for Early Detection and Prediction of Chronic Kidney Disease

Vijendra Singh [1,*] **, Vijayan K. Asari** [2] **and Rajkumar Rajasekaran** [3]

1 School of Computer Science, University of Petroleum and Energy Studies, Dehradun 248007, India
2 Electrical and Computer Engineering, University of Dayton, Dayton, OH 45469, USA; vasari1@udayton.edu
3 School of Computing Science and Engineering, Vellore Institute of Technology, Vellore 632014, India; rrajkumar@vit.ac.in
* Correspondence: vsingh.fet@gmail.com

Abstract: Diabetes and high blood pressure are the primary causes of Chronic Kidney Disease (CKD). Glomerular Filtration Rate (GFR) and kidney damage markers are used by researchers around the world to identify CKD as a condition that leads to reduced renal function over time. A person with CKD has a higher chance of dying young. Doctors face a difficult task in diagnosing the different diseases linked to CKD at an early stage in order to prevent the disease. This research presents a novel deep learning model for the early detection and prediction of CKD. This research objectives to create a deep neural network and compare its performance to that of other contemporary machine learning techniques. In tests, the average of the associated features was used to replace all missing values in the database. After that, the neural network's optimum parameters were fixed by establishing the parameters and running multiple trials. The foremost important features were selected by Recursive Feature Elimination (RFE). Hemoglobin, Specific Gravity, Serum Creatinine, Red Blood Cell Count, Albumin, Packed Cell Volume, and Hypertension were found as key features in the RFE. Selected features were passed to machine learning models for classification purposes. The proposed Deep neural model outperformed the other four classifiers (Support Vector Machine (SVM), *K*-Nearest Neighbor (*K*NN), Logistic regression, Random Forest, and Naive Bayes classifier) by achieving 100% accuracy. The proposed approach could be a useful tool for nephrologists in detecting CKD.

Keywords: chronic kidney disease; feature selection; recursive feature elimination; support vector machine; machine learning

1. Introduction

Chronic kidney disease is a disorder that occurs when a patient's kidney function deteriorates. As a result, their overall quality of life suffers. Chronic kidney disease affects one out of every 10 people worldwide (CKD). CKD is on the rise, and by 2040, it is expected to be the fifth leading cause of death worldwide [1]. It is one of the leading causes of high medical costs. In high-income nations, the cost of transplantation and dialysis accounts for 2% to 3% of the annual medical budget [2]. Most people with renal failure in low- and middle-income countries have insufficient access to life-saving dialysis and kidney transplants [3]. The number of kidney failure cases is expected to rise unexpectedly in developing countries such as China and India [4]. Chronic kidney failure makes to difficulties in removing extra fluids from the body blood. Advanced chronic kidney disease can cause dangerous levels of fluid, electrolytes, and wastes to build up in the body. It may lead to complications such as high blood pressure, anemia, weak bones, and nerve damage. The strongest indicator of renal function is the Glomerular Filtration Rate (GFR) [5]. Doctors also determine kidney disease through glomerular filtration rate (GFR). The criteria for defining CKD are a kidney damage for ≥ 3 months with or without decreased GFR or glomerular filtration rate (GFR) less than 60 mL/min/1.73 m^2 for ≥ 3 months with or

without kidney damage. GFR [6] is the most accurate predictor of kidney function for detecting different phases of CKD, with each phase indicating a more severe reduction in glomerular filtration rate.

The GFR [6] is also used to detect renal failure; if GFR < 15 mL/min, then it means kidney has failed or near to failure. This is the last (fifth) stage of chronic kidney disease. The diagnosis of CKD is a difficult task in medicine because it is based on a variety of symptoms. Clinical judgments are primarily dependent on physicians' expertise and experience during the examination of the patient's symptoms [7]. As the health care system evolves and new medicines become accessible, physicians are finding it increasingly challenging to stay up with current clinical practice changes [8].

The machine learning technique provides valid decision-making approaches for computer-assisted automatic disease identification [9]. Machine learning is being used to intelligently interpret available data and transform it into useful knowledge to increase the diagnostic process efficiency [1]. Machine learning is already being used to assess the state of the human body, analyze disease-related aspects, and diagnose a variety of disorders [10]. Heart disease has been diagnosed using models based on machine learning techniques [11]. Diabetes, heart disease, and retinopathy [12], acute renal injury [13], and cancer [14] were all diagnosed using models created by machine learning algorithms. Many researchers have used supervised algorithms, such as Random Forest [15], Fuzzy C Means [16], Naive Bayes [17], Support Vector Machine [18–21], Gradient Boosting [19–22], Logistic Regression [20] classifiers in detecting chronic kidney disease. Machine learning can also help to enhance the quality of medical data, reduce the frequency of hospital admissions, and save money on medical expenses. As a result, these models are more commonly utilized in diagnostic analytic research than other older approaches [23]. The only way to reduce chronic disease (CD) mortality is to diagnose it early and treat it effectively [24]. The feature extraction and classification processes in traditional machine learning involve two separate methods. As a result, typical machine learning approaches take a long time to compute. Because of this, the traditional technique is no longer viable for real-time diagnostic applications.

The ability of Artificial Neural Networks (ANN) to learn about fault tolerance, generalization, and the environment is becoming more widespread in the area of medical diagnostics.

In many cases, the Neural Networks (NN) method outperforms standard machine learning techniques. The resource learning architecture can be enhanced to boost its performance even more. The neural networks [25–29] models have been used for the detection of kidney disease. The majority of currently available CKD models have a low classification accuracy. As a result, this research introduces a novel model for Chronic Kidney Disease.

The main contributions of this paper are:

Deep neural networks have been proposed to detect and diagnose CKD.

Apply feature selection to improve efficiency and efficacy of deep neural network.

The computational accuracy of the proposed model is compared with existing classification methods from the literature.

Furthermore, the performance is evaluated through the various performance measures.

The following is a breakdown of the paper structure: The related works on machine learning approaches in the fields of CKD are presented in Section 2. Section 3 presents the proposed deep neural model for early detection of CKD. The results are discussed in Section 4, along with a full explanation. Section 5 wraps up the findings and looks ahead to the future.

2. Related Works

Machine learning models have been shown to be effective in predicting and diagnosing serious diseases. Early detection of chronic diseases, particularly the search for new treatments for chronic kidney disease, has gotten a lot of attention from doctors and researchers in recent years. Several recent research has demonstrated that machine

learning and deep learning models may be used to successfully diagnose chronic kidney disease (CKD). Table 1 presents a detailed comparison of machine learning methods for the diagnosis of Chronic Diseases from the existing literature.

Z. Chen, X.et al. proved the reliability of multivariate models in clinical practice risk assessment for patients with CKD [30]. The Chronic Renal Failure (CFR) data bank at UC Irvine was used in this investigation. In their comparison investigation, they used the *K*NN, SVM, and Soft independent modeling of class analogy. In comparison to the other two models, the SVM model processed noise within the data set better. In this comparison, the SVM accuracy was 99%. The author of [31] developed a decision-making tool for doctors to forecast the occurrence of CRF in patients. The authors employed *K*NN, Naive Bayes, LDA, random subspace, and tree-based classification techniques on the CRF data set from the UCI repository. The random subspace with the *K*NN classifier has a 94% accuracy rate, according to the researchers. The authors of another study [32] created decision support similar to [31]. The authors classified CRF using Artificial Neural Networks (ANN), Naive Bayes, and decision tree algorithms in this paper. The performance of various machine learning algorithms was examined on Jordan's Prince Hamza Hospital data set. The decision tree is predicted the most accurate when compared to two other approaches. Song et al. [22] created a gradient boosting-based prediction model to detect CKD using diabetes patient's EHR and billing data. The authors of [33] published a study on UCI CKD data sets that used SVM, decision trees, Nave Bayes, and *K*NN to detect CKD. The authors developed a ranking algorithm to choose features. With a score of 99.75, the decision tree outperformed three alternative machine learning methods. The authors of [34] presented a hierarchical multiclass classification technique for detecting chronic renal disease in an unbalanced data set.

As a baseline, the authors used naive Bayes, logistic regression, decision trees, and random forests classifiers. Within each patient, the proposed classification approach discovered severe cases. A chronic renal disease diagnosis system was proposed in [35] to diagnose CKD at an early stage. For preparing the data, the authors used the *K*-means technique. On processed data, the *K*NN, SVM, and Naive Bayes classification algorithms were used. Classification algorithms produced the greatest accuracy of 97.8%. Almasoud and Ward [36] reported a study on CKD that used logistic regression, SVM, random forest, and gradient boosting techniques. Four categorization techniques were applied to selected features. Gradient boosting has the highest accuracy of 99%. E M Senan et al. [37] recommended a study on early-stage CKD diagnosis. The RFE method was used to select characteristics from the CKD data set. The outcomes of the SVM, *K*NN, random forest and decision tree algorithms were compared.

Krishnamurthy S. et al. [38] developed various artificial intelligence models to predict Chronic Kidney Disease. The LightGBM model selected the most important features for CKD prediction: age, gout, diabetes mellitus, use of sulfonamides, and angiotensins. The convolutional neural networks achieved the best performance and the highest AUROC metric, 0.954, compared to other models. Mohamed Elhoseny et al. [19] presented an intelligent prediction system for Chronic Kidney Disease. The density-based Feature Selection method eliminates the irrelevant features and then passes selected features to the Ant Colony-based Optimization classifier to predict CKD. Singh and Jain [39] presented novel hybrid approach for diagnose CKD and achieved 92.5 % of prediction accuracy. An artificial neural network for CKD diagnosis was proposed by Neves et al. [25]. The diagnostic sensitivity values ranged from 93.1% to 94.9%, and the diagnostic specificity values ranged from 91.9% to 94.2% in this work.

Vasquez-Morales et al. [27] used large CKD data to generate a neural network classifier, and the model was 95% accurate. Makino et al. [28] used textual data to extract patients diagnoses and treatment information in order to forecast the course of diabetic kidney disease. Ren et al. [29] developed a predictive model for the identification of CKD from an Electronic Health Records (EHR) data set. This proposed model is based on a neural network framework that encodes and decodes the textual and numerical information

from EHR. Ma F. et al. [40] develop a deep neural network model to detect chronic renal disease. The presented model obtained the highest accuracy compared to ANN and SVM. Almansour et al. [41] devised a way for preventing CKD using machine learning. The SVM and ANN were among the machine learning classification algorithms used by the researchers. The results of the experiments revealed that ANN had a greater accuracy of 99.75% than SVM.

J. Qin et al. [42] presented a machine learning method for the early detection of CKD. They used logistic regression, random forest, SVM, naive Bayes classifier, *K*NN, and feed-forward neural network to develop their models. The most accurate classification model was random forest, which had a 99.75% accuracy rate. Z. Segal et al. [43] presented a machine learning technique based on an ensemble tree (XGBoost) for the early diagnosis of renal illness. The presented model was compared against Random Forest, CatBoost, Regression with Regularization. The proposed model showed better performance in all matrices, including c-statistics 0.93, sensitivity 0.715, and specificity 0.958. Khamparia et al. [44] developed a deep learning model for early detection of CKD, in which features were selected from multimedia data using a stacked autoencoder model. The authors used A SoftMax classifier to predict the final class. It was observed that the proposed model achieved the highest performance in comparison to conventional classification techniques on the UCI CKD data set.

Polat, H. et al. [45] presented a study on the role of effective feature selection methods in the accurate prediction of CKD. In this paper, wrapper and filter feature selection approaches were used to select the dimension of the Chronic Kidney Disease data set. The selected features are then passed to Support Vector Machine to classify Chronic Kidney Disease for diagnosis purposes. The experimental results presented that Support Vector Machine generates better results on selected features by the Best First search method with filtered subset evaluator. SVM achieved an accuracy rate (98.5%) in comparison to features selected by other wrapper and filter methods. Ebiaredoh-Mienye Sarah A. et al. [46] presented a robust model for the prediction of CKD that integrates an enhanced sparse autoencoder (SAE) and Softmax regression. In this proposed model, autoencoders achieved sparsity by penalizing the weights. The Softmax regression was optimized for the classification task; therefore, the proposed model achieved excellent performance. The proposed model obtained an accuracy of 98% on the chronic kidney disease (CKD) data set. The proposed model achieved comparable performance with other existing methods. Zhiyong Pang et al. [47] proposed a fully automated computer-aided diagnosis system to classify malignant and benign masses using breast magnetic resonance imaging. The texture features were selected by integration of support vector machine with ReliefF feature selection method. This system achieved an accuracy of 92.3%. Chen et al. [21] presented a model in which Hepatitis was diagnosed with a hybrid method that integrates a Fisher discriminatory analysis algorithm and an SVM classifier. As a result of comparing the proposed method with the existing methods, the hybrid method outperforms the other methods, and the highest classification accuracy of 96.77% is achieved. The authors presented a breast cancer diagnosis model in this study [48]. The selected features by sequential forward selection and the backward selection methods are passed to Artificial Neural Networks to classify breast cancer. SBSP + NN achieved the highest accuracy of 98.75%.

Table 1. Comparative Accuracy analysis of the diagnosis of Chronic Diseases from literature.

Chronic Diseases Diagnosis	Model	Accuracy Achieved (%)	Reference
Kidney Renal Failure	Artificial Neural Networks	91.9%–94.2%	[25]
Diabetic Kidney Disease	Convolutional Model	71%	[28]
Chronic Kidney Disease	Neural Network Classifier	95%	[27]
Breast Cancer	SBSP + NN	98.57%	[48]

Table 1. *Cont.*

Chronic Diseases Diagnosis	Model	Accuracy Achieved (%)	Reference
Hepatitis Disease	FDA and SVM	96.77%	[21]
Breast Cancer	SVM + ReliefF	92.3%	[47]
Chronic Kidney Disease	*KNN*, SVM	99%	[30]
Chronic Renal Failure	Fisher Discriminatory Analysis and SVM	96.7%	[21]
Chronic Renal Failure	*KNN*, Naive Bayes, LDA, Random Subspace and Tree-Based Decision	94%	[31]
Chronic Kidney Disease	SVM, *KNN*, and Naïve Bayes Decision tree,	99.7%	[33]
Chronic Kidney Disease	Logistic Regression, Decision Tree, Naïve Bayes, and Random Forests	93%	[34]
Chronic Kidney Disease	*KNN*, SVM, and Naïve Bayes	97.8%	[35]
Chronic Kidney Disease	SVM, *KNN*, and decision tree	99.1%	[37]
Chronic Kidney Disease	Convolutional Neural Networks	95.7%	[38]
Chronic Kidney Disease	SVM, Random Forest, and Gradient Boosting	99%	[36]
Chronic Kidney Disease	Logistic regression, *KNN*, SVM, Random Forest, Naive Bayes and ANN	99.7%	[42]
Chronic Kidney Disease	XGBoost	95.8%	[43]
Chronic Kidney Disease	SVM	98.5%	[45]
Chronic Kidney Disease	Softmax Regression	98%	[46]

3. Materials and Methods

3.1. Data Set Description

The University of California Irvine (UCI) Repository was used to gather CKD data. There are 400 patient records in the data set, and some values are missing. It comprises 24 clinical qualities that emerge in the prognosis of chronic kidney disease, with one class attribute serving as a result of the patient's presence of chronic renal failure being predicted. There are two types of values in the expected feature diagnostic: "ckd" and "notckd." The data set contains 250 values of the "ckd" class (62.5%) and 150 values of the "notckd" class (37.5%). The characteristics of the UCI CKD data collection are listed in Table 2.

Table 2. Characteristics of the UCI CKD data.

Features	Specification	Value
AGE	AGE (IN YEARS)	0–90
AL	ALBUMIN	0–5
ANE	ANAEMIA	NO, YES
APPET	APPETITE	POOR, GOOD
BA	BACTERIA	PRESENT, NOTPRESENT
BGR	BLOOD GLUCOSE RANDOM	0–490
BP	BLOOD PRESSURE	0–180
BU	BLOOD UREA	0–391
CAD	CORONARY ARTERY DISEASE	NO, YES
CLASS	CLASS	NOTCKD, CKD
DM	DIABETES MELLITUS	NO, YES
HEMO	HAEMOGLOBIN	0–17.8
HTN	HYPERTENSION	NO, YES
PC	PUS CELL	NORMAL, ABNORMAL
PCC	PUS CELL CLUMPS	PRESENT, NOTPRESENT

Table 2. *Cont.*

Features	Specification	Value
PCV	PACKED CELL VOLUME	0–54
PE	PEDAL EDEMA	NO, YES
POT	POTASSIUM	0–47
RBC	RED BLOOD CELLS	NORMAL, ABNORMAL
RC	RED BLOOD CELL COUNT	0–8
SC	SERUM CREATININE	0–76
SG	SPECIFIC GRAVITY	0–1.025
SOD	SODIUM	0–163
SU	SUGAR	0–5
WC	WHITE BLOOD CELL COUNT	0–26,400

3.2. Data Processing

The estimation of missing values and the removal of noise such as outliers, as well as the normalization and validation of unbalanced data, were all part of the preprocessing stages. When assessing a patient, some measurements could be missing or incomplete.

3.2.1. Handling Missing Values

There are 158 completed cases in the data collection, with the remainder missing. Ignoring records is the simplest technique to deal with missing values; however, this is not practical for small data sets. The data set is examined during the data preparation process to see whether any attribute values are missing. The missing values for numerical features were estimated using the statistical technique of mean imputation. The mode technique was used to replace the missing values of nominal features.

3.2.2. Categorical Data Encoding

Because most machine learning algorithms only accept numeric values as input, category values must be encoded into numerical values. The binary values "0" and "1" are used to represent the characteristics of categories such as "no" and "yes".

3.2.3. Data Transformation

Data transformation is the process of transforming numbers on the same scale so that one variable does not dominate the others. Otherwise, learning algorithms perceive larger values as higher and smaller values as lower, regardless of the unit of weight. Data transformations alter the values in a data set so that they can be processed further [49]. To improve the accuracy of machine learning models, this research employs a data normalization technique. It converts data between the −1 and +1 ranges. The converted data has a standard deviation of 1 and a mean of 0.

The standardization formula is given below:

$$w = \frac{(x - \overline{x})}{\sigma} \tag{1}$$

where,
w = Standardized score
x = Observed value
$\overline{x}$ = Mean
σ = Standard deviation

3.2.4. Outlier Detection

Outliers are observation points that are isolated from the rest of the data. An outlier could be caused by measurement variability or signal an error in the experiment. An outlier can distort and mislead the learning process of the machine learning algorithm. It leads to longer training times, less accuracy in the model, and ultimately to poorer results. This

paper uses the Interquartile Range (IQR) [49] based approach to remove outliers before transferring data to the learning algorithm.

3.3. Feature Selection

Recursive Feature Elimination (RFE) removes features recursively, building a model based on other features [50]. It applies greedy search to find the most efficient subset of features. Use model accuracy to determine which features are most appropriate for predicting a feature. It develops models iteratively, determining the best or worst feature for each iteration. The traits are then classified based on the sequence in which they were removed. If the data set contains N functions, recursive feature elimination will eagerly search for a combination of 2N features in the worst-case scenario.

3.4. Classifiers

3.4.1. Support Vector Machine

The SVM constructs a separation hyperplane that splits the labeled data into classes and determines whether a new data value belongs above or below the line. There may be several hyperplanes, and the one with the largest margin between data points is chosen. Figure 1 shows the maximum hyperplanes and maximum margin of the support vector machine. The equation of hyperplane that separates two classes is given by:
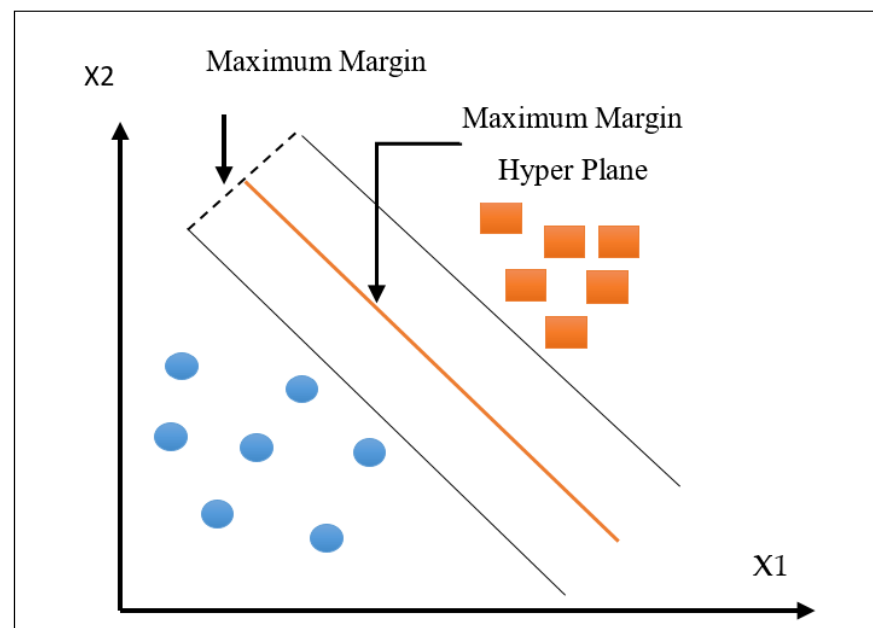
$$D(x) = w_0 + w_1 a_1 + w_2 a_2 \tag{2}$$

Figure 1. Support Vector Machine.

However, the equation of the maximum-margin hyperplane can be written

$$x = b + \sum_i \propto_i y_i a(i) \times a \tag{3}$$

Here, i is the support vector, and y_i is the training instance $a(i)$ class value. The learning algorithm determines the numeric value b and α_i, respectively.

3.4.2. *K*-Nearest Neighbor

The *K*NN algorithm recognizes similarities between new and previous data points and categorizes fresh test points into existing related groups. The *K*NN method is a slow learning algorithm since it is not parametric. This means that instead of learning from

the training data set, it should be secured. It uses *K* to categorize the data. The distance between the new location and the saved training point was determined using the Euclidean distance. Figure 2 depicts *K*-Nearest neighbor classification based on *K* values.

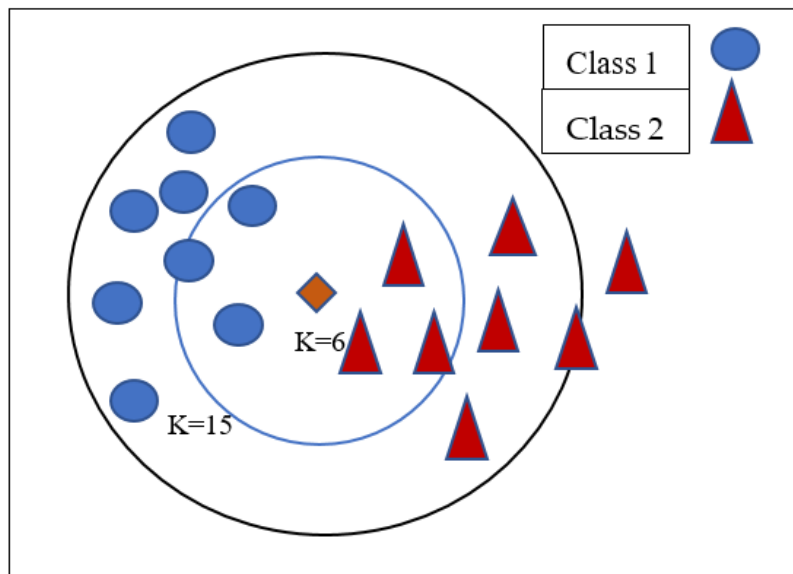$$d_{ij} = \sum_{t=1}^{n} \left(x_{it}^{test} - x_{jt}^{train} \right)^2 \tag{4}$$

Figure 2. *K*-Nearest Neighbor.

*K*NN algorithm searches *t* training data set with minimum distance to the testing set.

3.4.3. Decision Tree Classifier

Decision trees are a nonparametric method of supervised learning [51]. This is a classified structured tree that defines the characteristics of a data set. It represents internal rules for decision-making through internal nodes and tree branches. It has two types of nodes, the decision, and the leaf nodes. The decision nodes take some decisions, and the outcomes of such decisions are leaf nodes. A decision tree has presented in Figure 3.
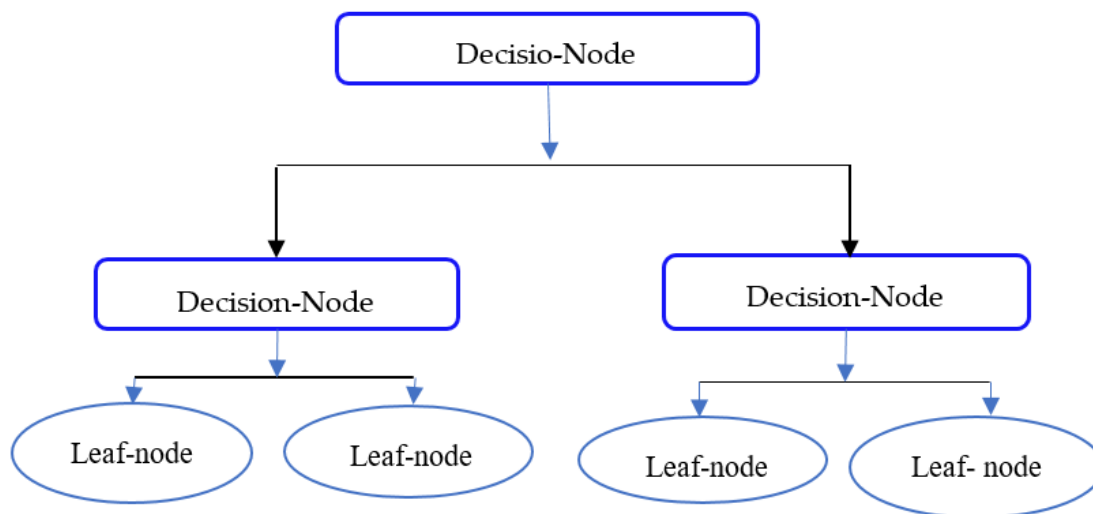

Figure 3. Decision trees.

3.4.4. Random Forest Classifier

The random forest algorithm is based on ensemble learning, improving the model's performance, and solving complex problems by combining several classifiers. A classifier named after the algorithm that contains multiple decision trees averaged over a database subset to improve predictions. In the forecasting process, it does not rely on a single decision tree, and the random forest algorithm creates forecasts from each decision tree that predicts the conclusion based on the majority of decision votes. The usage of several trees decreases the possibility of the model overfitting. To predict the classes in the database, the algorithm includes many decision trees, some of which can predict the proper outcome while others cannot. As a result, there are two assumptions regarding the prediction's accuracy. To forecast a more accurate outcome than an estimate, the algorithm must first include the actual value of the feature variable. Second, there must be an extremely low correlation between the forecasts for each tree. As a result, there are two requirements for high forecast accuracy. Figure 4 shows a Random Forest Classifier.
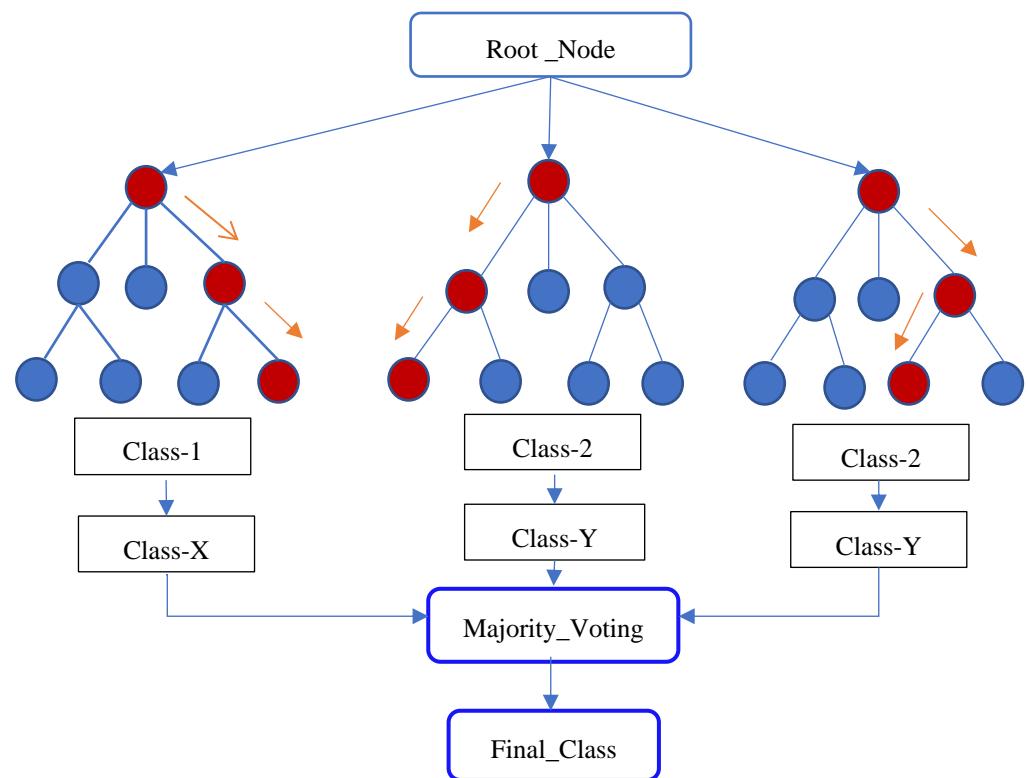

Figure 4. Random Forest.

3.5. Model Development

Figure 5 depicts the model's framework. Preprocessing, model hyper tuning, and classification are the three phases of the proposed model. Because the data set may contain noise and redundant values, the preparation step is the most important.

This phase applied different methods such as handling missing values, categorical data encoding, data transformation, removing outliers and extreme values, and feature selection. The CKD data set is separated into training and testing data sets after being preprocessed. Only a few features are selected using Recursive Feature Elimination out of a total of 24 features in this study. The RFE algorithm evaluates each feature's value based on its significance, which helps to lower the method's processing complexity. Finally, redundant and unrelated characteristics are filtered away. The learning model is then fed with the most important characteristics. Figure 6 shows the pseudo-code for the proposed

methodology. Initially, a method was introduced to prepare and standardize the data in the data set. The processed data is further passed for processing.
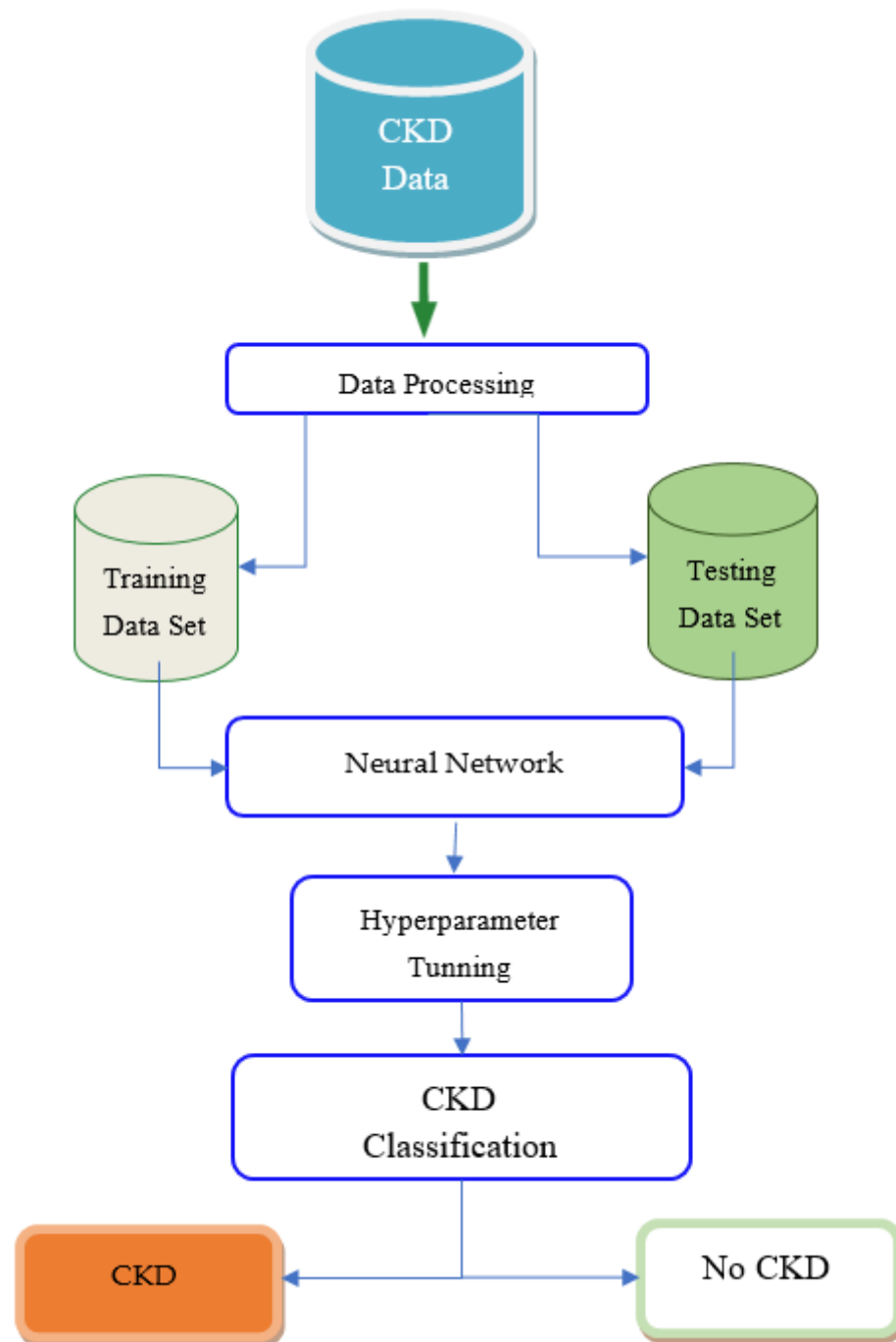

Figure 5. A framework of the proposed model.

Pseudo-code

BEGIN

1. Importing a CKD data set

2. Fill in the blanks and eliminate outliers

3. Convert text into numeric values

4. Adjust the data scale

5. Make use of the RFF 'Feature Selection.'

6. FOR FEATURE SELECTION METHODS:

7. Choose the most important features.

8. Make a list of the most important features.

9. Count the number of times each feature appears.

END FOR

10. Provide the proposed model with a list of features.

11. Set the model hyperparameters

12. Assign different validation scores to the scoring.

13. Train using the proposed model

14. Sort data into CKD and non-CKD categories.

15. Validation score

END

Figure 6. Pseudo-Code of the proposed model.

There are 12 layers in the proposed model architecture: an input layer, five dense layers, five drop layers, and an output dense classifier layer. In Figure 7, the layered architecture's exact specifications are depicted. Each dense layer is connected directly in a feed-forward method in this architecture. The layer is built in such a way that the outputs of its activation maps are handed on to all following levels as input. A dropout layer is placed between two dense layers in this model, with drop rates of 0.5, 0.4, 0.3, 0.2, and 0.1. Figure 7 presents the layered architecture of the proposed model.

The CNN model has several hyperparameters that need to be optimized. The optimal hyperparameters selection process is experimental; however, it is time-consuming and difficult. Adam [52,53] optimizer initiates hyperparameters with smaller parameters during the training phase.

Adam uses adaptive assessment to determine individual learning rates for various hyperparameter grades ranging from first to second-order gradients. Stochastic Gradient Optimization (SGD) [54] is less efficient than Adam. It necessitates minimal learning time and memory. The classification performance is enhanced by the CNN correct activation function. Neural network's standard activation functions are sigmoid, tan, Rectified Linear Unit (ReLU) [55], Exponential Linear Unit (ELU) [56], and Self-Normalized Linear Unit (SELU) [57]. This paper tested the different activation functions on the CKD data set and selected the preferred one in all the models.
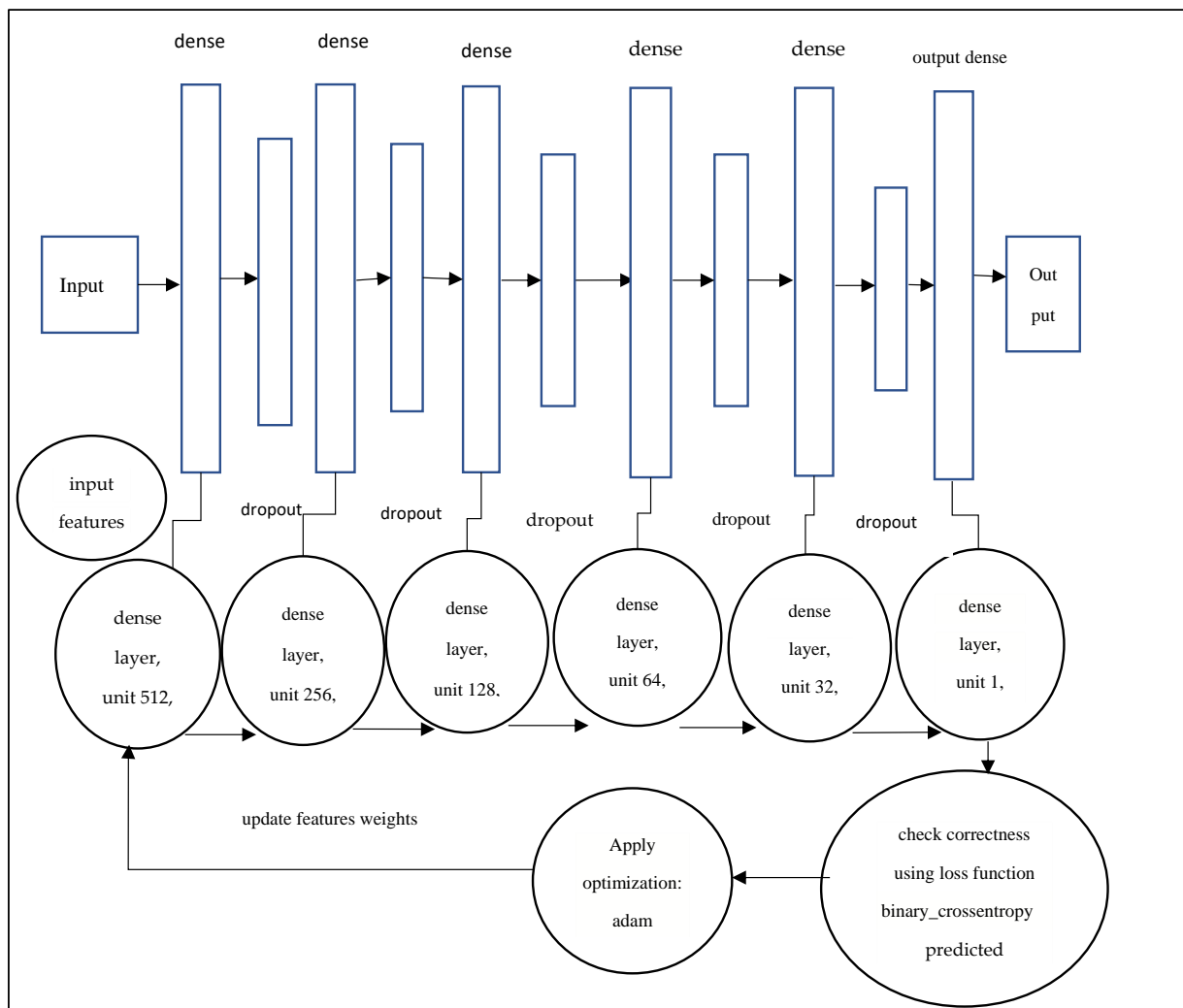
Figure 7. Layers architecture of the proposed model deep neural network.

4. Results and Discussion

4.1. Experiment Setup

The proposed model was created using data from a variety of situations. The configuration of the system of the developing model is shown in Table 3.

Table 3. Experimental setup details.

Resource	Specification
Processor	Intel Core i5 Gen7
Random access memory	16 GB
Graphics processing unit	4 GB
Language	Python

4.2. Evaluation Parameters

The proposed model accuracy was calculated by making the CKD class value positive and the notCKD class value negative. The confusion matrix was utilized to evaluate the performance by using True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN) [58]. According to TP, CKD samples have been accurately categorized. The findings of the FN test show that CKD samples were misclassified. The notCKD samples were not accurately identified, as indicated by a false-positive result (FP). True negative (TN) samples have been accurately categorized as not CKD.

4.2.1. Accuracy

It refers to the proportion of correct guesses to total predictions. Accuracy can be described as the ability to accurately predict the outcome of a situation.

$$\text{Accuracy} = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \tag{5}$$

4.2.2. Recall

The recall calculates the proportion of accurately predicted positive observations to the total number of observations in the class, as shown in the following equation.

$$\text{Recall} = \frac{T_P}{T_P + F_N} \tag{6}$$

4.2.3. Specificity

The specificity estimates the number of well-scored negative patterns. The higher the specificity value, the more negative the classifier. It can be defined as:

$$\text{Specificity} = \frac{T_N}{T_N + F_P} \tag{7}$$

4.2.4. Precision

As stated in the equation below, this metric represents the proportion of accurately predicted positive observations to total predictive positive observations.

$$\text{Precision} = \frac{T_P}{T_P + F_P} \tag{8}$$

4.2.5. F-Measure

Precision and Recall are weighted averaged in the F-measure [58]. False positives and false negatives are part of the process. F-measure is a term that is defined as

$$\text{F-Measure} = \frac{\text{Two} \times (\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})} \tag{9}$$

The F-Measure values lie from 0 to 1.

4.3. Comparative Analysis of Results

The findings of the proposed model are presented in this section. The CKD data sets are split into 75% training and 25% test data sets. The hyperparameter settings for the proposed model are shown in Table 4. The confusion matrices are shown in Figure 8. It demonstrates that the suggested model correctly identified all genuine positive and true negative events. The CKD class reports recall, precision, sensitivity, F1 score, and accuracy.

Table 4. Hyper-parameter settings.

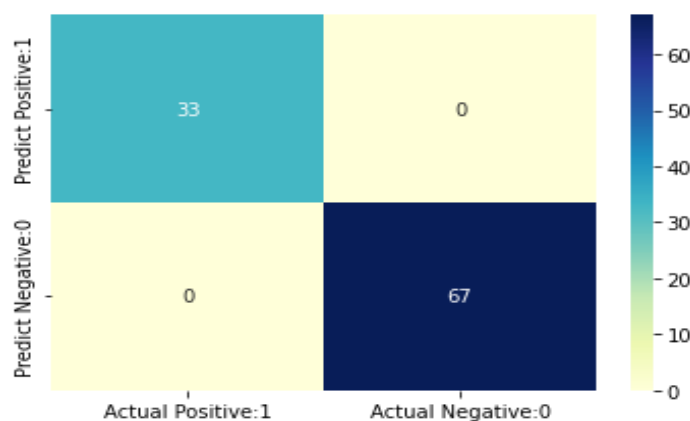| Hyper-Parameter	Setting
Epochs	850
Batch size	15
Dropout rate	0.5 to 0.1
Activation Function	relu
Activation output layer	sigmoid
Optimizer	Adam
Loss	binary_crossentropy

Figure 8. Confusion matrices of the Proposed model.

The proposed model is compared with other classifier algorithms, including logistic regression, *K*NN, SVM, Decision tree, and Random forest. No parameter adjustments were made for these algorithms to show the improved performance of the proposed model. Therefore, the default value for a parameter was used in scikit-learn. All models are evaluated using the F1-score. Tables 5 and 6 showed experimental results when the proposed model was tested on CKD data sets. In contrast,

Table 5. Comparative analysis of the proposed model with existing classification techniques on CKD data set.

Method	Accuracy	Recall	Precision	F-Measure
Logistic Regression	0.99	1.0	0.98	0.99
K-Nearest Neighbor	0.92	0.88	0.98	0.92
Naïve bayes	0.95	0.92	1.00	0.95
Support Vector Machines	0.92	0.87	0.96	0.92
Decision Tree	0.97	0.95	1.00	0.97
Proposed Model	1.00	1.00	1.00	1.00

Table 6. Comparative analysis of the proposed model with existing models from the literature on the UCI data set.

Authors	Model	Accuracy (%)
Elhoseny et al. [19]	Ant Colony-based Optimization Classifier	95
Vasquez-Morales et al. [27]	Neural network	95
M Senan et al. [37]	*K*NN	98.33
Krishnamurthy et al. [38]	Convolutional Neural Networks	95.4
Polat, H et al. [45]	Support Vector Machine	98.5
Sarah A. et al. [46]	SAE and Softmax Regression	98
Proposed Model	Deep Neural Network	100

Figures 9 and 10 depict accuracy graphs comparing the performance of existing classification algorithms to the proposed approach for chronic kidney disease prediction.
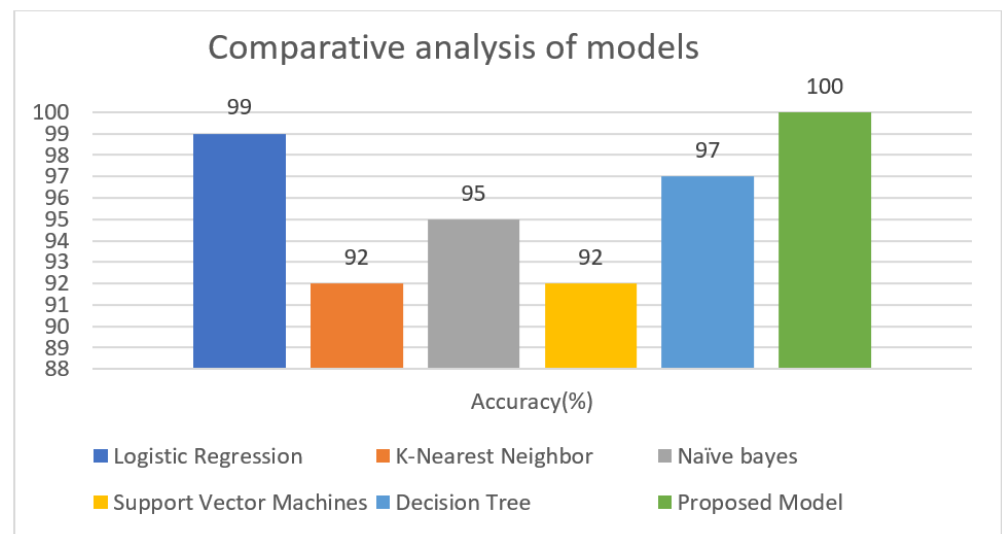
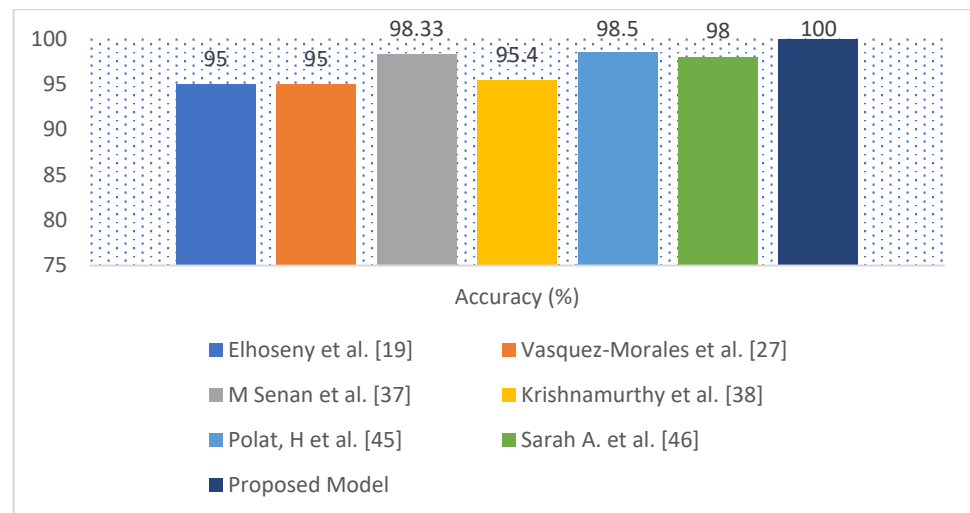**Figure 9.** Accuracy graphical representation for the UCI CKD data set.

Figure 10. Accuracy graphical representation for the UCI CKD data set.

The accuracy of *K*NN, SVM, Naïve Bayes, Decision tree, logistic regression, and the proposed model is 92%, 92%, 95%, 97%, 99% and 100%, respectively. The proposed model was found to be the most accurate, with a 100% accuracy rate. Because it optimally identified positive samples as 250 samples (TP) and all 150 samples as negative samples, the suggested model appropriately classifies all positive and negative samples (TN). True Positive samples were graded 99%, 92%, 95%, 92%, and 97% by Logistic Regression, *K*NN, Nave Bayes, SVM, and Decision Tree, respectively, with a margin of error of 1%, 8%, 5%, 8%, and 3%, respectively. The results of all five classifiers are shown in Table 5.

The proposed model outperforms the other classifiers by scoring 100% on all measures. The F1-score, accuracy, precision, and recall of the Logistic regression were all 99%, 99%, 100%, and 98%, respectively. Then Decision Tree obtained an F1-score, Accuracy, Precision, and Recall of 97%, 97%, 95%, and 100%, respectively. The Naïve Bayes F1-score, Accuracy, Precision, and Recall values were 95%, 95%, 92%, and 100%, respectively. The F1-score, Accuracy, Precision, and Recall values of Naïve Bayes were 92%, 92%, 88%, and 98%, respectively. The Support Vector Machines classifier performed the lowest with F1-score, Accuracy, Precision, and Recall values of 92%, 92%, 87%, and 96%, respectively.

Table 6 compares the proposed model to several recent scholarly studies, such as Ant Colony-based Optimization Classifier by Elhoseny et al. [19], Neural network by

Vasquez-Morales et al. [27], *K*NN by M Senan et al. [37], Convolutional Neural Networks by Krishnamurthy et al. [38], SVM by Polat, H. et al. [45], and SAE and Softmax Regression proposed by Sarah, A. et al. [46]. The proposed model has obtained an accuracy of 100%, while the exiting works obtained the accuracy from 85% to 98.5%. Finally, it should be noted that the proposed model is more efficient than existing classification methods.

4.4. Feature Importance from RFE

This section of the paper presents the most important feature selected by the RFE algorithm based on their ranking. The figure shows the chosen features and their importance during the classification of CKD disease. The most critical risk factors are Hemoglobin, Serum Creatinine, Specific Gravity, Packed Cell Volume, Red Blood Cell Count, Hypertension, and Albumin, as presented in Table 7. The nephrologists should focus on these risk factors while diagnosing CKD disease patients. Figure 11 shows feature selected by RFE with their importance.

Table 7. The most critical risk factors from CKD data.

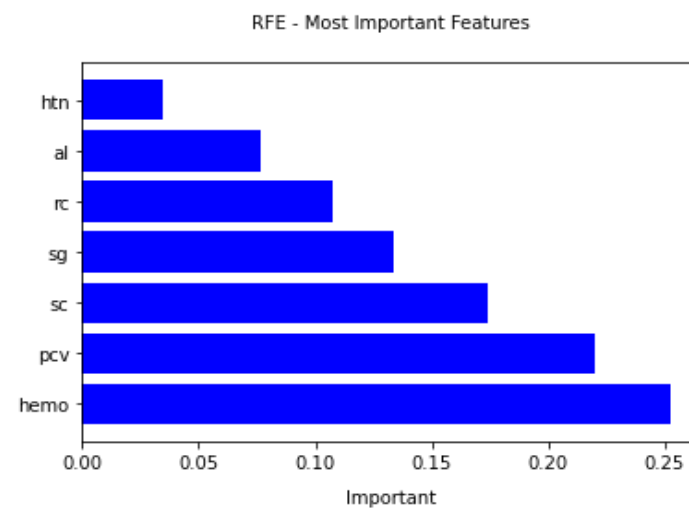| Risk Factor Name |
| --- |
| Hemoglobin |
| Serum Creatinine |
| Red Blood Cell Count |
| Packed Cell Volume |
| Albumin |
| Specific Gravity |
| Hypertension |

Figure 11. Important features selected by RFE.

4.5. Receiver Operating Characteristic (ROC)/Area under Curve (AUC)

The bottom of the square and the ROC curve define the area of the AUC. AUC scores closer to 1 indicate good performance, whereas AUC scores closer to 0.50 indicate poor performance. Figures 12–17 shows the ROC/AUC curve of the proposed model, logistic regression, Decision tree, SVM, *K*NN, and Naïve Bayes respectively. The proposed model achieved the highest AUC score value 1.0.
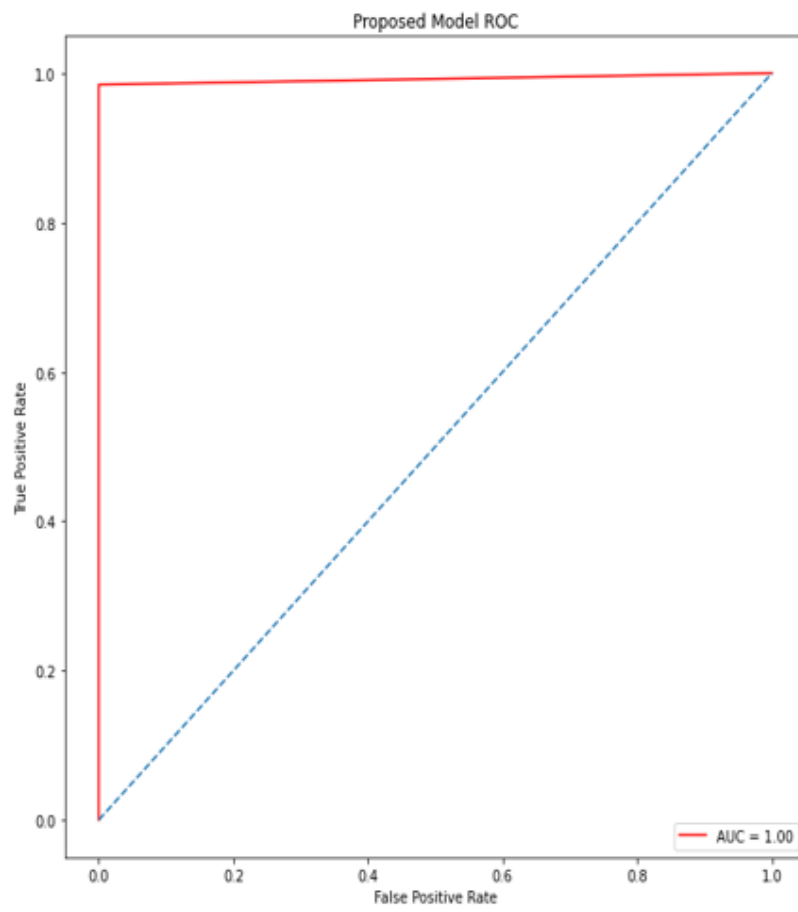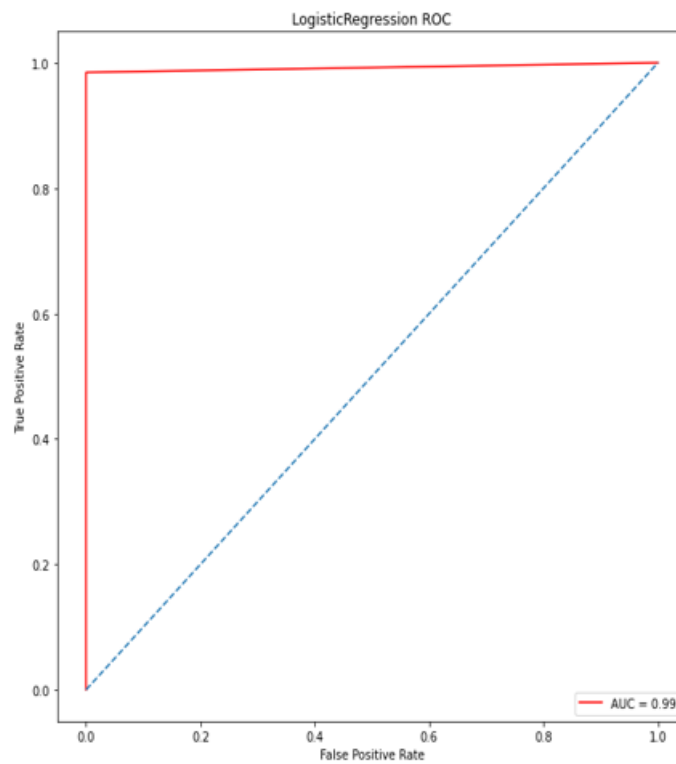
Figure 12. ROC/AUC of Proposed model.
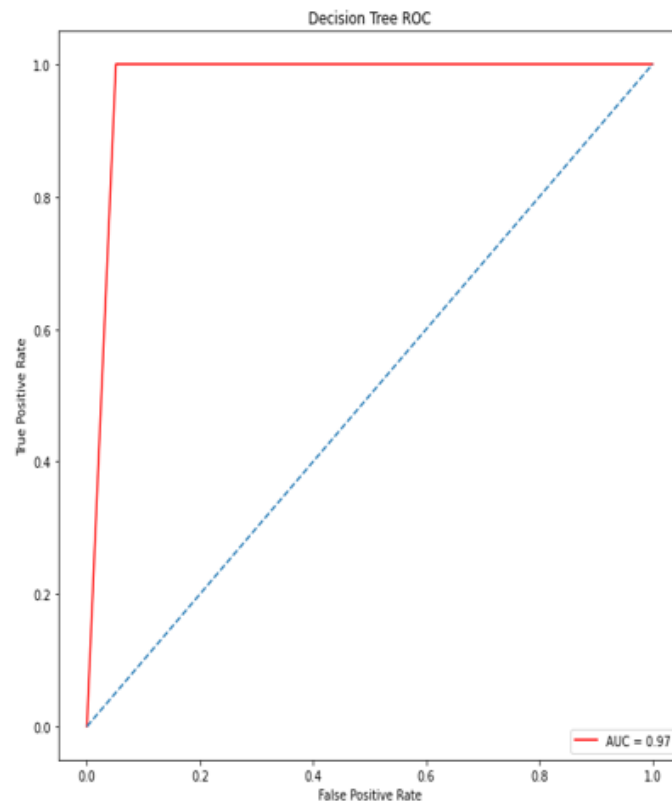

Figure 13. ROC/AUC of Logistic Regression.

Figure 14. ROC/AUC of Decision tree.
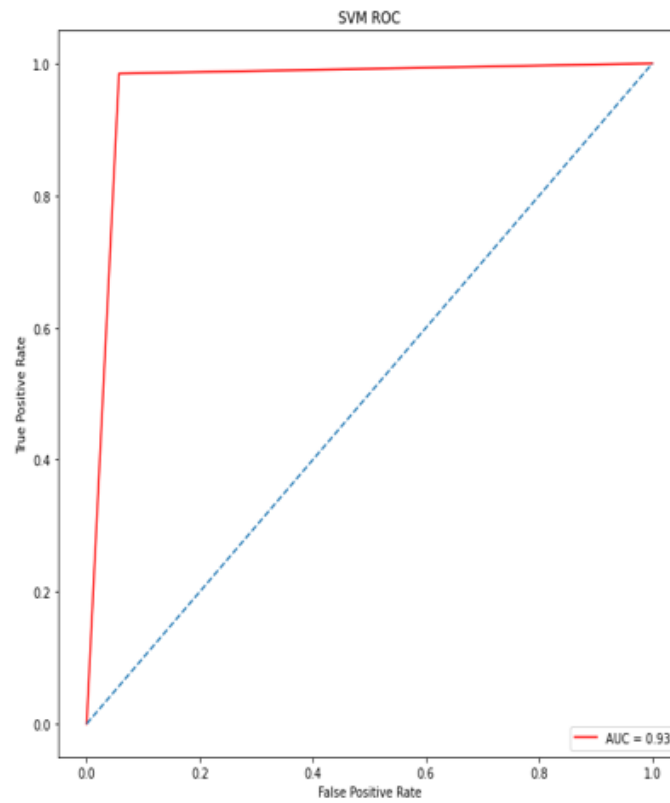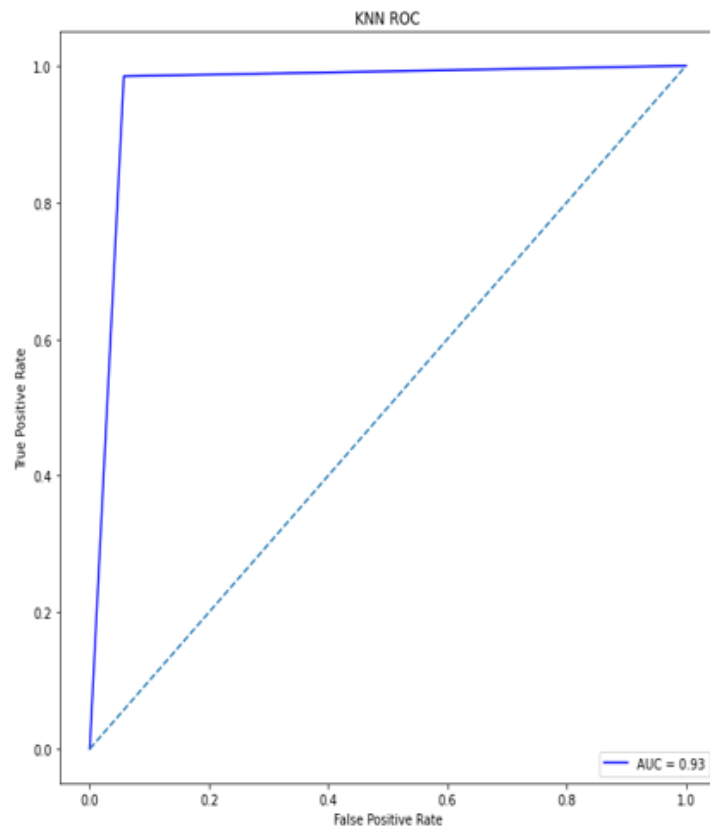

Figure 15. ROC/AUC of SVM.

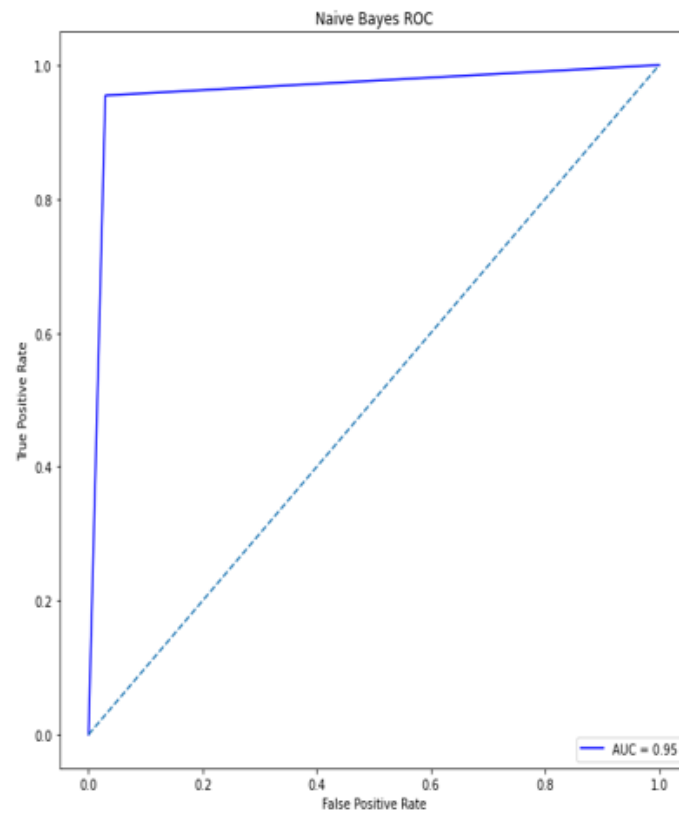**Figure 16.** ROC/AUC of *K*NN.

Figure 17. ROC/AUC of Naïve bayes.

5. Conclusions and Future Work

A deep learning model for the early diagnosis of chronic disease is presented in this work. In this research, the authors looked at the Recursive Feature Elimination approach to identify which features are the most important for prediction. The most essential CKD features are packed red blood cell count, albumin, cell volume, serum creatinine, specific gravity, hemoglobin, and hypertension. Classification algorithms are fed with a set of features. Different metrics, including classification accuracy, recall, precision, and f-measure, are used to estimate the comparative analysis. The proposed deep neural model outperformed the other five classifiers (Support Vector Machine (SVM), *K*-Nearest Neighbor (*K*NN), Logistic regression, Random Forest, and Naive Bayes classifier) by achieving 100% accuracy. The accuracy of *K*NN, SVM, Naïve Bayes, Decision tree, Random Forest, logistic regression is 92%, 92%, 95%, 97%, and 99%, respectively.

The performance of the proposed model compared with several recent scholarly studies, such as Ant Colony-based Optimization Classifier by Elhoseny et al. [19], Neural network by Vasquez-Morales et al. [27], *K*NN by M Senan et al. [37], Convolutional Neural Networks by Krishnamurthy et al. [38], SVM by Polat, H. et al. [45], and SAE and Softmax Regression proposed by Sarah A. et al. [46]. The exiting works obtained the accuracy from 85% to 98.5%, while the proposed model has obtained an accuracy of 100%. The proposed approach could be a useful tool for nephrologists in detecting CKD.

The limitation of the proposed model was that it had been tested on small data sets. To improve the model performance, significant volumes of increasingly sophisticated and representative CKD data will be collected in the future to detect disease severity. The clinical data to be collected from pathologist's experts. The performance of the proposed model will be evaluated on a large clinical data set based on acid-base parameters, hyperparathyroidism, inorganic phosphorus concentration, and night urination in the future. Additionally, new features will be applied to get a broader perspective on the informative parameters related to CKD disease to test the prediction accuracy.

Author Contributions: Conceptualization, V.S.; Methodology, V.S.; Software, V.S.; Validation, V.S. and V.K.A.; Formal Analysis, V.S.; Data Curation, R.R. and V.S.; Writing—Original Draft Preparation, V.S.; Writing—Review and Editing, V.S., V.K.A. and R.R.; visualization, V.S.; Supervision, V.S., V.K.A. and R.R.; project administration, V.S. All authors have read and agreed to the published version of the manuscript.

Informed Consent Statement: No need for this research work.

Data Availability Statement: Data were collected from UCI Machine Repository, CA, USA (http://archive.ics.uci.edu/ml (accessed on 18 June 2021). Source code and data are available at https://github.com/vsingh-fet/Deep_Neural (accessed on 28 December 2021).

References

1. Foreman, K.J.; Marquez, N.; Dolgert, A.; Fukutaki, K.; Fullman, N.; McGaughey, M.; Pletcher, M.A.; Smith, A.E.; Tang, K.; Yuan, C.-W.; et al. Forecasting life expectancy, years of life lost and all-Cause and cause-Specific mortality for 250 causes of death: Reference and alternative scenarios for 2016-40 for 195 countries and territories. *Lancet* **2018**, *392*, 2052–2090. [CrossRef]
2. Vanholder, R.; Annemans, L.; Brown, E.; Gansevoort, R.; Gout-Zwart, J.J.; Lameire, N.; Morton, R.L.; Oberbauer, R.; Postma, M.J.; Tonelli, M.; et al. Reducing the costs of chronic kidney disease while delivering quality health care: A call to action. *Nat. Rev. Nephrol.* **2017**, *13*, 393–409. [CrossRef] [PubMed]
3. 2020 Wkd Theme. Available online: https://www.worldkidneyday.org/2020-campaign/2020-wkd-theme/ (accessed on 20 July 2021).
4. Jha, V.; Garcia-Garcia, G.; Iseki, K.; Li, Z.; Naicker, S.; Plattner, B.; Saran, R.; Wang, A.Y.-M.; Yang, C.-W. Chronic kidney disease: Global dimension and perspectives. *Lancet* **2013**, *382*, 260–272. [CrossRef]
5. Nataional Kidney Foundation. 2020. Available online: https://www.kidney.org/kidneydisease/global-facts-about-kidney-disease (accessed on 20 July 2021).

6. Levin, A.S.; Bilous, R.W.; Coresh, J. Chapter 1: Definition and classification of CKD. *Kidney Int. Suppl.* **2013**, *3*, 19–62.
7. Chen, T.K.; Knicely, D.H.; Grams, M.E. Chronic kidney disease diagnosis and management: A review. *JAMA* **2019**, *322*, 1294–1304. [CrossRef]
8. Meesad, P.; Yen, G. Combined numerical and linguistic knowledge representation and its application to medical diagnosis. *IEEE Trans. Syst. Man Cybern.-Part A Syst. Hum.* **2003**, *33*, 206–222. [CrossRef]
9. Gürbüz, E.; Kılıç, E. A new adaptive support vector machine for diagnosis of diseases. *Expert Syst.* **2014**, *31*, 389–397. [CrossRef]
10. Mahyoub, M.; Randles, M.; Baker, T.; Yang, P. Comparison Analysis of Machine Learning Algorithms to Rank Alzheimer's Disease Risk Factors by Importance. In Proceedings of the 2018 11th International Conference on Developments in eSystems Engineering (DeSE), Cambridge, UK, 2–5 September 2018; pp. 1–11.
11. Masetic, Z.; Subasi, A. Congestive heart failure detection using random forest classifier. *Comput. Methods Programs Biomed.* **2016**, *130*, 54–64. [CrossRef]
12. Gao, Z.; Li, J.; Guo, J.; Chen, Y.; Yi, Z.; Zhong, J. Diagnosis of Diabetic Retinopathy Using Deep Neural Networks. *IEEE Access* **2019**, *7*, 3360–3370. [CrossRef]
13. Park, N.; Kang, E.; Park, M.; Lee, H.; Kang, H.-G.; Yoon, H.-J.; Kang, U. Predicting acute kidney injury in cancer patients using heterogeneous and irregular data. *PLoS ONE* **2018**, *13*, e0199839. [CrossRef]
14. Patrício, M.; Pereira, J.; Crisóstomo, J.; Matafome, P.; Gomes, M.; Seiça, R.; Caramelo, F. Using Resistin, glucose, age and BMI to predict the presence of breast cancer. *BMC Cancer* **2018**, *18*, 29. [CrossRef]
15. Ilyas, H.; Ali, S.; Ponum, M.; Hasan, O.; Mahmood, M.T.; Iftikhar, M.; Malik, M.H. Chronic kidney disease diagnosis using decision tree algorithms. *BMC Nephrol.* **2021**, *22*, 273. [CrossRef]
16. Ahmed, K.A.; Aljahdali, S.; Hussain, S.N. Comparative prediction performance with support vector machine and random forest classification techniques. *Int. J. Comput. Appl.* **2013**, *69*, 12–16.
17. Drall, S.; Drall, G.S.; Singh, S.; Naib, B.B. Chronic kidney disease prediction using machine learning: A new approach. *Int. J. Manag. Technol. Eng.* **2018**, *8*, 278–287.
18. Balija, R.; Sriraam, N.; Geetha, M. Classification of non-chronic and chronic kidney disease using SVM neural networks. *Int. J. Eng. Technol.* **2017**, *7*, 191–194.
19. Elhoseny, M.; Shankar, K.; Uthayakumar, J. Intelligent diagnostic prediction and classification system for chronic kidney disease. *Sci. Rep.* **2019**, *9*, 9583. [CrossRef] [PubMed]
20. Fisher, M.A.; Taylor, G.W. A Prediction Model for Chronic Kidney Disease Includes Periodontal Disease. *J. Periodontol.* **2009**, *80*, 16–23. [CrossRef]
21. Chen, H.; Liu, D.-Y.; Yang, B.; Liu, J.; Wang, G. A new hybrid method based on local fisher discriminant analysis and support vector machines for hepatitis disease diagnosis. *Expert Syst. Appl.* **2011**, *38*, 11796–11803. [CrossRef]
22. Song, X.; Waitman, L.R.; Yu, A.S.; Robbins, D.C.; Hu, Y.; Liu, M. Longitudinal risk prediction of chronic kidney disease in diabetic patients using temporal-enhanced gradient boosting machine: Retrospective cohort study. *JMIR Med. Inf.* **2020**, *8*, e15510. [CrossRef]
23. Napolitano, G.; Marshall, A.; Hamilton, P.; Gavin, A.T. Machine learning classification of surgical pathology reports and chunk recognition for information extraction noise reduction. *Artif. Intell. Med.* **2016**, *70*, 77–83. [CrossRef]
24. Eslamizadeh, G.; Barati, R. Heart murmur detection based on wavelet transformation and a synergy between artificial neural network and modified neighbor annealing methods. *Artif. Intell. Med.* **2017**, *78*, 23–40. [CrossRef] [PubMed]
25. Neves, J.; Martins, M.R.; Vilhena, J.; Neves, J.; Gomes, S.; Abelha, A.; Machado, J.; Vicente, H. A soft computing approach to kidney diseases evaluation. *J. Med. Syst.* **2015**, *39*, 131. [CrossRef]
26. Di Noia, T.; Ostuni, V.C.; Pesce, F.; Binetti, G.; Naso, D.; Schena, F.P.; di Sciascio, E. An end stage kidney disease predictor based on an artificial neural networks ensemble. *Expert Syst. Appl.* **2013**, *40*, 4438–4445. [CrossRef]
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. [CrossRef]
28. Makino, M.; Yoshimoto, R.; Ono, M.; Itoko, T.; Katsuki, T.; Koseki, A.; Kudo, M.; Haida, K.; Kuroda, J.; Yanagiya, R.; et al. Artificial intelligence predicts the progression of diabetic kidney disease using big data machine learning. *Sci. Rep.* **2019**, *9*, 11862. [CrossRef] [PubMed]
29. Ren, Y.; Fei, H.; Liang, X.; Ji, D.; Cheng, M. A hybrid neural network model for predicting kidney disease in hypertension patients based on electronic health records. *BMC Med. Inf. Decis. Mak.* **2019**, *19*, 131–138. [CrossRef] [PubMed]
30. Chen, Z.; Zhang, X.; Zhang, Z. Clinical risk assessment of patients with chronic kidney disease by using clinical data and multivariate models. *Int. Urol. Nephrol.* **2016**, *48*, 2069–2075. [CrossRef]
31. Al-Hyari, A.Y.; Al-Taee, A.M.; Al-Taee, M.A. Clinical decision support system for diagnosis and management of Chronic Renal Failure. In Proceedings of the 2013 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT), Amman, Jordan, 3–5 December 2013; pp. 1–6.
32. Ani, R.; Sasi, G.; Sankar, U.R.; Deepa, O.S. Decision support system for diagnosis and prediction of chronic renal failure using random subspace classification. In Proceedings of the 2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Jaipur, India, 21–24 September 2016; pp. 1287–1292.

33. Tazin, N.; Sabab, S.A.; Chowdhury, M.T. Diagnosis of Chronic Kidney Disease using effective classification and feature selection technique. In Proceedings of the 2016 International Conference on Medical Engineering, Health Informatics and Technology (MediTec), Dhaka, Bangladesh, 17–18 December 2016; pp. 1–6.
34. Bhattacharya, M.; Jurkovitz, C.; Shatkay, H. Chronic Kidney Disease stratification using office visit records: Handling data imbalance via hierarchical meta-classification. *BMC Med. Inf. Decis. Mak.* **2018**, *18*, 125. [CrossRef]
35. Akben, S. Early stage chronic kidney disease diagnosis by applying data mining methods to urinalysis, blood analysis and disease history. *IRBM* **2018**, *39*, 353–358. [CrossRef]
36. Almasoud, M.; Ward, T.E. Detection of chronic kidney disease using machine learning algorithms with least number of predictors. *Int. J. Soft Comput. Appl.* **2019**, *10*, 89–96. [CrossRef]
37. 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.; Jadhav, M.E.; Alzahrani, M.Y. Diagnosis of Chronic Kidney Disease Using Effective Classification Algorithms and Recursive Feature Elimination Techniques. *J. Healthc. Eng.* **2021**, *2021*, 1004767. [CrossRef] [PubMed]
38. Krishnamurthy, S.; Ks, K.; Dovgan, E.; Luštrek, M.; Piletič, 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. [CrossRef] [PubMed]
39. Singh, V.; Jain, D. A Hybrid Parallel Classification Model for the Diagnosis of Chronic Kidney Disease. *Int. J. Interact. Multimed. Artif. Intell.* **2021**. [CrossRef]
40. Ma, F.; Sun, T.; Liu, L.; Jing, H. Detection and diagnosis of chronic kidney disease using deep learning-based heterogeneous modified artificial neural network. *Future Gener. Comput. Syst.* **2020**, *111*, 17–26. [CrossRef]
41. Almansour, N.; 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. *Comput. Biol. Med.* **2019**, *109*, 101–111. [CrossRef]
42. Qin, J.; Chen, L.; Liu, Y.; Liu, C.; Feng, C.; Chen, B. A Machine Learning Methodology for Diagnosing Chronic Kidney Disease. *IEEE Access* **2019**, *8*, 20991–21002. [CrossRef]
43. Segal, Z.; Kalifa, D.; Radinsky, K.; Ehrenberg, B.; Elad, G.; Maor, G.; Lewis, M.; Tibi, M.; Korn, L.; Koren, G. Machine learning algorithm for early detection of end-stage renal disease. *BMC Nephrol.* **2020**, *21*, 518. [CrossRef]
44. 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. *Multimed. Tools Appl.* **2020**, *79*, 35425–35440. [CrossRef]
45. Polat, H.; Mehr, H.D.; Cetin, A. Diagnosis of chronic kidney disease based on support vector machine by feature selection methods. *J. Med. Syst.* **2017**, *41*, 55. [CrossRef] [PubMed]
46. Ebiaredoh-Mienye, S.A.; Esenogho, E.; Swart, T.G. Integrating Enhanced Sparse Autoencoder-Based Artificial Neural Network Technique and Softmax Regression for Medical Diagnosis. *Electronics* **2020**, *9*, 1963. [CrossRef]
47. Pang, Z.; Zhu, D.; Chen, D.; Li, L.; Shao, Y. A Computer-Aided Diagnosis System for Dynamic Contrast-Enhanced MR Images Based on Level Set Segmentation and ReliefF Feature Selection. *Comput. Math. Methods Med.* **2015**, *2015*, 450531. [CrossRef]
48. Uzer, M.S.; Inan, O.; Yılmaz, N. A hybrid breast cancer detection system via neural network and feature selection based on SBS, SFS and PCA. *Neural Comput. Appl.* **2013**, *23*, 719–728. [CrossRef]
49. Jiawei, H.; Micheline, K.; Jian, P.S. *Data Mining: Concepts and Techniques*, 3rd ed.; Morgan Kaufmann Publishers: San Francisco, CA, USA, 2012.
50. Saeys, Y.; Inza, I.; Larrañaga, P. A review of feature selection techniques in bioinformatics. *Bioinformatics* **2007**, *23*, 2507–2517. [CrossRef]
51. Podgorelec, V.; Kokol, P.; Stiglic, B.; Rozman, I. Decision trees: An overview and their use in medicine. *J. Med. Syst.* **2002**, *26*, 445–463. [CrossRef]
52. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
53. Reddi, S.; Zaheer, M.; Sachan, D.; Kale, S.; Kumar, S. Adaptive methods for nonconvex optimization. In Proceeding of the 32nd Conference on Neural Information Processing Systems (NIPS 2018), Montroal, QC, Canada, 3–8 December 2018.
54. Orr, G.B.; Müller, K.R. *Neural Networks: Tricks of the Trade*; Springer: Berlin/Heidelberg, Germany, 2003.
55. Nair, V.; Hinton, G.E. Rectified linear units improve restricted boltzmann machines. In Proceedings of the 27th International Conference on Machine Learning, Haifa, Israel, 21–24 June 2010; pp. 1–8.
56. Clevert, D.A.; Unterthiner, T.; Hochreiter, S. Fast and accurate deep network learning by exponential linear units (elus). *arXiv* **2015**, arXiv:1511.07289.
57. Klambauer, G.; Unterthiner, T.; Mayr, A.; Hochreiter, S. Self-normalizing neural networks. In Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; pp. 972–981.
58. Powers, D.M. Evaluation: From precision, recall and f-measure to roc, informedness, markedness and correlation. *J. Mach. Learn. Technol.* **2011**, *2*, 37–63.