Research Article https://doi.org/10.65606/41/2026/79

# An Optimized Deep Belief Neural Network Approach for Chronic Kidney Disease Prediction Using GWO

**Ali Iftikhar Butt [1], Muhammad Hamza [2*]**

[1] Punjab University College of Information Technology, Lahore, Pakistan

[2] College of Science and Technology, Lahore, Pakistan

*Corresponding Author: Muhammad Hamza Author Email: Hamza112@gmail.com

**Received:** October 19, 2025 **Revised:** March 11, 2026 **Accepted:** June 21, 2026



**Abstract:** In today's healthcare system, accurate classification of medical data is crucial for effective disease diagnosis and prediction. Machine learning (ML) and deep learning (DL) models are extensively utilized for these tasks, yet they often face challenges related to computational complexity and processing efficiency. This research introduces a novel model for diagnosing chronic kidney disease (CKD) that integrates advanced data preprocessing, feature selection, and classification techniques. Missing data is addressed using Multiple Imputation by Chained Equations (MICE) to maintain dataset integrity, followed by feature selection using Grey Wolf Optimization (GWO) to enhance prediction accuracy. Finally, a Deep Belief Neural Network (DBNN) performs the classification. Model effectiveness is evaluated through sensitivity, accuracy, and specificity, demonstrating superior performance compared to existing methods.

**Keywords:** Grey Wolf Optimization (GWO), Chronic Kidney Disease (CKD), Feature Selection, Deep Belief Neural Networks (DBNNs)

## 1. Introduction

CKD is a long-term, progressive disease characterized by a steady decline in kidney function, leading to the accumulation of waste products in the body. This buildup can result in severe complications such as bone disease, anemia, and cardiovascular issues, significantly impacting patient quality of life [1]. For advanced cases of CKD, end-stage renal disease (ESRD) may develop, necessitating either dialysis or a kidney transplant for survival [2]. Approximately 10% of the global population suffers from CKD, which has become a leading cause of mortality due to the challenges associated with early detection [3, 4]. Traditional diagnostic methods, including urine and blood tests, are commonly employed; however, they often fail to identify CKD in its early stages, delaying treatment and accelerating disease progression [5]. The recent advances in machine learning (ML) technology have shown the enormous potential of applying it in the medical diagnostics, especially when it comes to detecting diseases at the earliest stages [6]. As research done by different scientists proves, ML technologies have the potential to detect subtle patterns and correlations in the large amounts of data, which helps in early diagnosis and prediction of such chronic conditions as CKD [7]. They can analyze the patient history and biochemical data and classify the risk factors of the development of CKD and make predictions about its appearance [8]. However, the use of ML faces problems in terms of the computational complexity, training time, and high dimensionality of medical data [9].

In order to solve the aforementioned problems, this paper offers a hybrid approach based on using Grey Wolf Optimization (GWO) method for selecting the best features and Deep Belief Neural Network (DBNN) for classification. By filtering out the unimportant features, GWO decreases the computational demands of CKD dataset, making the work of the machine more efficient. The selected features will be used by DBNN, which will classify the input and give high diagnostic accuracy without high computational burden [10]. The proposed method is assessed based on the performance metrics such as accuracy, sensitivity, and specificity.

## 2. Literature Review

Prediction of chronic diseases like CKD is getting more importance nowadays and advanced computational techniques like machine learning (ML) and deep learning (DL) can be helpful for achieving better diagnosis and efficiency in this regard [11]. In this connection, various researchers have tried different feature selection and classification algorithms along with optimization techniques in order to make better prediction models in relation to CKD.

### 2.1. Machine learning Techniques for CKD Prediction

Machine learning has largely been applied to predict CKD disease showing good results due to its ability to deal with complex datasets. Researchers such as [4] have designed a model of neural network using 10 fold cross validation technique which was superior to traditional machine learning models like decision trees and SVM in terms of accuracy in predicting CKD disease. Tekale [11] have also used SVM to classify CKD disease according to biochemical and demographic information with high accuracy although at higher computational cost. Moreover, researchers have also explored ensemble techniques like random forest (RF) and gradient boosting to overcome noise issue of medical datasets. Saha [12] have compared RF and Naïve bayes (NB) machine learning methods for CKD classification and concluded that these ensemble methods have better accuracy and robustness as compared to other techniques while these models require higher computational power.

### 2.2. Feature Selection Techniques in CKD Prediction

Feature selection plays an important role in improving predictive models through identifying attributes that are relevant to the model, resulting in lower dimensionality and overfitting of the model. Metaheuristic algorithms have gained popularity as one way of achieving this because they are able to use natural processes to determine optimal features. This technique has worked well for CKD datasets using algorithms such as Particle Swarm Optimization (PSO) and Ant Colony Optimization (ACO). For instance, according to Kamel [13], feature selection using PSO enhanced classsification accuracy while at the same time reducing computational cost; a strategy that has been successfully used by [6] who combined PSO and Artificial Neural Network (ANN). The Grey Wolf Optimization (GWO) is a relatively new metaheuristic algorithm that mimics the social hunting process of grey wolves. It has been proven effective in medical predictions including the selection of features used in predicting diabetes; a task in which it was able to reduce dimensionality yet improve predictability.

### 2.3. Deep Learning in CKD Prediction

Deep learning models, especially those with hierarchical structures, are increasingly used in CKD diagnosis due to their capacity to learn complex, high-dimensional patterns. Deep Belief Networks (DBNN) and Convolutional Neural Networks (CNN) are particularly effective for medical image and data analysis. Gupta [2] applied a DBNN model to CKD datasets, achieving 98.5% accuracy due to the model's ability to capture intricate data representations (Gupta et al., 2020). CNNs, while primarily applied to image processing, have been adapted for structured medical data. Li et al. (2019) developed a hybrid CNN model that used layers to hierarchically process CKD attributes, achieving improvements in accuracy and sensitivity. Similarly, Munguía-Siu [14] Humayun [18] used a CNN-RNN hybrid for disease prediction, which further illustrated the advantages of combining CNN with feature extraction methods in complex datasets.

### 2.4. Hybrid Approaches Combining Optimization and Deep Learning

Combining optimization techniques with DL models has shown great promise in enhancing diagnostic accuracy while reducing computational overhead. Wang [15] implemented a hybrid AlexNet-based transfer learning model, integrating data augmentation and PSO-based feature selection for alcoholism detection, demonstrating that transfer learning combined with feature selection can improve diagnostic precision. For CKD diagnosis, Abdelaziz [8] proposed a model using cloud-based frameworks for integrating Linear Regression with Neural Networks, achieving efficient data

processing in a cloud environment. Researchers implemented a similar approach for cancer prediction, combining GWO and DBNNs, achieving significant improvements in classification accuracy and computation times [16]. These hybrid approaches highlight the possibilities of integrating optimization techniques with deep learning models to generate accurate and computationally efficient solutions.

In this regard, this research work is an advancement that introduces a GWO-DBNN hybrid model specifically for the purpose of CKD prediction. In this model, GWO works to enhance feature selection through dimensionality reduction and computational efficiency, whereas the DBNN works on improving diagnostic performance. This model is an efficient alternative to conventional machine learning and deep learning models in diagnosing CKD in real time [1].

The work has been painstakingly structured to improve readers' understanding. The first sections of the publication include a summary of the CKD detection introduction. A thorough review of the current methods for diagnosing CKD is given in Section II. The recommended method for identifying CKD is examined in part III. The fourth part provides a thorough analysis of the algorithm's performance. A comprehensive overview that includes our conclusion is provided in Section V.

## 3. Methodology

### 3.1 Data Set

The recommended model is based on the UCI Dataset [19]. It comprises 25 characteristics, 11 numeric, 14 nominal, and 1 class attribute, representing data samples from 400 individuals, 250 of whom have CKD and 150 have not. Table 1 portrays the dataset.

This work offers a fresh strategy to address the problems associated with patient illness prediction. Our main goal is to use advanced medical data categorisation procedures to advance illness forecast accuracy. In particular, we distillate on carefully choosing features to maximise the classification performance for the CKD data samples. We will go into depth about the thorough technique we used in our research in the parts that follow.

### 3.2 Data Preprocessing Phase: Handling Missing Values

Handling missing data is essential for reliable CKD classification. MICE is a robust technique that iteratively imputes missing values by generating regression models for each incomplete attribute, based on the observed values in other attributes. Each missing entry x_ij for attribute j and case i is estimated as:

$$x\_ij=f\_j\ (x\_(i,-j)\ )+\in$$

Where f_i is the predictive model for j, (x_(i,-j) ) represents other observed values in i, and $\epsilon$ captures uncertainty. This iterative process produces multiple datasets, preserving natural data variability and improving the robustness of classification models. MICE is preferred here as it maintains statistical integrity and is suitable for datasets with interdependent, complex missing data structures.

Feature selection eliminates irrelevant and noisy dataset features to retain the most pertinent information, crucial for effective classification. Datasets often carry superfluous attributes that can harm classification outcomes. Hence, feature selection is employed to reduce dataset dimensionality by removing undesirable attributes. In this study, the Grey Wolf Optimization (GWO) algorithm was employed, treating each feature as a Grey Wolf, to filter and select optimal features for classification [23]. This process yields a subset with the most informative attributes. Feature selection is vital for improving learning and classification processes, especially when overfitting risk grows with excessive features. The GWO algorithm is key to determining an optimal feature set for enhanced performance.

**Table 1.** Depiction of the dataset used

| ID | Attribute Symbols | Attribute Category | Scale |
|---|---|---|---|
| f1 | Age of patient | Numerical | Years |
| f2 | Blood pressure | Numerical | Mm/Hg |
| f3 | Specific gravity | Nominal | 0.005,1.010,1.015,1.020,1.025 |
| f4 | Albumin | Nominal | 0,1,2,3,4,5 |
| f5 | Sugar | Nominal | 0,1,2,3,4,5 |
| f6 | Red blood cells | Nominal | Normal, Abnormal |
| f7 | Pus cells | Nominal | Normal, Abnormal |
| f8 | Pus cells clumps | Nominal | Present , not present |
| f9 | Bacteria | Nominal | Present, not present |
| f10 | Blood glucose Random | Numerical | Mgs/dl |
| f11 | Blood urea | Numerical | Mgs/dl |
| f12 | Serum creatinine | Numerical | - |
| f13 | Sodium | Numerical | mEq/L |
| f14 | Potassium | Numerical | - |
| f15 | Hemoglobin | Numerical | Gms |
| f16 | Packed cell volume | Numerical | - |
| f17 | White blood cell count | Numerical | Cells/cum |
| f18 | Red blood cell count | Numerical | Millions/cm |
| f19 | Hypertension | Nominal | Yes , no |
| f20 | Diabetes mellitus | Nominal | Yes, no |
| f21 | Coronary artery disease | Nominal | Yes, no |
| f22 | Appetite | Nominal | Good, poor |
| f23 | Pedal edema | Nominal | Yes, no |
| f24 | Anemia | Nominal | Yes, no |
| f25 | Class | Nominal | CKD,NOTCKD |

### 3.3 GWO algorithm

Grey wolf predation behaviour is modelled by the evolutionary computing approach known as the GWO algorithm [17]. A wolf pack has five to twelve intelligent members. When they hunt, they form small groups and work together rather than acting alone. Alpha (α), beta (β), delta (δ), and omega (ω) are the four categories into which the grey wolves are classified within the pack based on the dominant hierarchy. Leaders known as alphas are in charge of choosing when to hunt, rest, and take action. Helping alphas make decisions and with other tasks is the main duty of betas. Although they may regulate omegas, deltas must submit to alphas and betas. Omegas must submit to all other alpha wolves. There are two periods in the predation model of grey wolves. First, the prey is surrounded by the wolves, as explained by:

$$\vec{X}(t+1) = \vec{X}_p(t) + \vec{A} \cdot \vec{D} \qquad (1)$$

Where t signify the iterations, $\vec{X}$ designates the position of the wolf, $\vec{X}_p$ designates the position of the target, $\vec{D}$ is described as

$$\vec{D} = \left|\vec{C} \cdot \vec{X}_p(t) - \vec{X}(t)\right| \qquad (2)$$

The $\vec{A}$ and $\vec{C}$ are described as below:

$$\vec{A} = 2a \cdot \vec{r_1} - a \qquad (3)$$

$$\vec{C} = 2 \cdot \vec{r_2}, \qquad (4)$$

Where $\overrightarrow{r_1}$ and $\overrightarrow{r_2}$ represent the random vectors inside [0, 1]. The GWO method ranks alphas as the best candidate results, betas as the second-finest candidate results, and deltas as the third-best. Alphas, betas, and deltas have the most acquaintance about the position of the food. While they reach their best places, other searchers are obliged to upgrade their placements as well. Hence, to continue assaulting the prey, wolves must update their location by the following equation as below:

$$\vec{X}(t+1) = \frac{\overrightarrow{X_1}+\overrightarrow{X_2}+\overrightarrow{X_3}}{3} \quad (5)$$

$$\begin{aligned} \overrightarrow{X_1} &= \left|\overrightarrow{X_\alpha} - \overrightarrow{A_1}\cdot\overrightarrow{D_\alpha}\right|, \\ \overrightarrow{X_2} &= \left|\overrightarrow{X_\beta} - \overrightarrow{A_2}\cdot\overrightarrow{D_\beta}\right|, \\ \overrightarrow{X_3} &= \left|\overrightarrow{X_\delta} - \overrightarrow{A_3}\cdot\overrightarrow{D_\delta}\right|, \end{aligned} \quad (6)$$

where $\overrightarrow{X_1}$, $\overrightarrow{X_2}$ and $\overrightarrow{X_3}$ are designed by the following equations;

Where $\overrightarrow{X_\alpha}$, $\overrightarrow{X_\beta}$, $\overrightarrow{X_\delta}$ represent the optimum first three finest solutions for every iterations, and $\overrightarrow{A_1}$, $\overrightarrow{A_2}$, $\overrightarrow{A_3}$ could be designed by Equation. (5). The $\overrightarrow{D_\alpha}$, $\overrightarrow{D_\beta}$, $\overrightarrow{D_\delta}$ can be attained by the following equations:

$$\begin{aligned} \overrightarrow{D_\alpha} &= \left|\overrightarrow{C_1}\cdot\overrightarrow{X_\alpha} - \vec{X}\right| \\ \overrightarrow{D_\beta} &= \left|\overrightarrow{C_2}\cdot\overrightarrow{X_\beta} - \vec{X}\right| \\ \overrightarrow{D_\delta} &= \left|\overrightarrow{C_3}\cdot\overrightarrow{X_\delta} - \vec{X}\right|, \end{aligned} \quad (7)$$

Where $\overrightarrow{C_1}$, $\overrightarrow{C_2}$ and $\overrightarrow{C_3}$ are intended by Equation. (6). these two stages are recurrent till the wolves seize the prey.

### 3.4 CKD data classification

The nominated features from BGWO are fixed to CKD classification level. In this stage, we focus on training a Deep Neural Network (DNN) for CKD classification. This DNN approach, also known as an ANN, comprises multiple hidden layer and output units. Our training process involves two critical phases: Pre-training: During this period, the neural network is initialized by utilizing a generative DBN. Based on hidden unit activations, the DBN enables the network to construct meaningful initializations. It entails of an input layer with input characteristics, many hidden layers, and an output layer with one unit per class. A Restricted Boltzmann Machine (RBM) is used to tackle the primary problem in DNN training.

RBM ignores class labels in the data and is first used as an unsupervised training method. Every data sample is processed using RBM, following certain guidelines. The activation vector produced by the RBM's hidden units mirrors the dependencies and patterns in the data. A probabilistic distribution of data attributes is what this vector looks like. The vector is then run through the RBM again in reverse, resulting in a "confabulation" of the data. A crucial phase of feature learning is essentially learnt by the RBM, which learns to encode and decode unique data characteristics. These altered characteristics are ready for subsequent steps, such as classification fine-tuning in a DNN.

$$C\left(u_i = \frac{1}{y}\right) = \sigma(b_i + \sum W_{ji} y_j \quad (8)$$

$$C\left(y_i = \frac{1}{u}\right) = \sigma(c_i + \sum W_{ji} u_j \quad (9)$$

In the context of performing stochastic sharpest scaling in log probability with the training numbers, this model illuminates an exceptionally simple learning principle.

$$\Delta_{wij}\alpha < u_i y_i >_{data} - < u_i y_i >_{reconstruction} \quad (10)$$

Below is Pseudocode of GWO algorithm for feature selection:

Input: N_iter: Maximum iterations for optimization, N_wolves: Number of grey wolves (features) in the population nd Dim: Dimensionality of the feature space (number of features)
Output: Optimal subset of selected features
Initialize:

- Initialize the positions of N_wolves in the feature space at random in [0, 1]
- Set alpha ($\alpha$), beta ($\beta$), delta ($\delta$) to the best three wolves (solutions)
- Set a, A, C parameters according to Equations (3) & (4)

While stopping criteria not met (iteration < N_iter):
For each wolf (i) in the pack:

- Evaluate the fitness (objective function) based on feature subset quality (e.g., classification accuracy)
- Update the positions of wolves based on $\alpha$, $\beta$, and $\delta$ by employing Equations (6) &(7)

Update parameter 'a' (linearly reductions from 2 to 0 with iterations)
For each wolf (i) in the pack:

- Calculate coefficient vectors A and C:

For each wolf (i)

- Calculate the distance from prey according to Equation (2).
- Update positions based on the positions of $\alpha$, $\beta$, $\delta$ using Equations (5),(6) &(7)
- Update wolf's new position according to Equation (5).
- Convert wolf positions to binary using a transfer function (sigmoid or similar):

If sigmoid(X_i) > threshold, select feature; else, discard feature

- Identify new alpha (finest solution), beta (second-best), and delta (third-best) wolves based on the updated fitness values.

End While
Output: The positions of the best wolf ($\alpha$) represent the optimal subset of selected features.

Once an RBM is trained, it can be "loaded" to construct a multilayer system. When a new RBM is loaded, the input layer becomes a prepared vector, and unit values for the layers in the already trained RBM are determined using the current biases and weights. The last layer of the previously trained layers interacts with the input for the new RBM. This updated RBM is trained using this process, and this entire procedure continues until a specified stopping criterion is met.

After the network has undergone pretraining as the RBM, the pre-training weights are further refined through backpropagation, aiming to enhance the overall accuracy of the system. This final stage, known as the fine-tuning stage, utilizes the standard backpropagation procedure.

## 4. Results and Discussion

### 4.1 Material (Software and Hardware Used

The existence of CKD in the examined dataset was identified using GWO feature selection and the DBN classifier, which was conducted in MATLAB Simulink on an Intel Core i7, 2.60 GHz CPU with 8GB RAM. Performance parameters such as Accuracy, Recall, Precision, and F-Measure, and compared to projected results.

### 4.2. Model Evaluation

The dataset is split at random into two sections, with the first containing 70% of the total data collected used to train the algorithm. The second component is utilized for testing and contains the dataset's reminder (30%). Six performance gauges are utilized to assess and validate the suggested model. These measurements are Accuracy, Recall, Precision, and F-Measure.

Next equations show how these measures computed

$$\text{Accuracy} = \frac{t_p + t_n}{t_p + f_p + f_n + t_n} \qquad (11)$$

$$\text{Sensitivity} = \frac{t_p}{t_p + f_n} \quad (12)$$

$$\text{Specificity} = \frac{t_n}{t_n + f_p} \quad (13)$$

$$\text{Kappa} = \frac{p_o - p_a}{1 - p_a} \quad (14)$$

$$F - \text{Score} = \frac{2 \times \text{Recall} \times \text{Precision}}{(\text{Recall} + \text{Precision})} \quad (15)$$

In this study, 15 optimal features were selected from the CKD dataset using the BGWO algorithm, as detailed in Table 4. Table 5 compares the performance of various feature selection methods, including GWO, GA, PSO and BPSO algorithms. The accuracy results, presented in Figure 1, show that for feature sets ranging from 10 to 60, the GWO algorithm consistently outperformed other methods like PSO, Binary PSO (BPSO) and GA in terms of accuracy on the CKD dataset.

**Table 2.** Acquired optimal features by GWO algorithm

| Attributes Label | Attributes Description |
|---|---|
| Age | Patient's age |
| Blood pressure | Indication of heart rate |
| Specific gravity | The ratio of urine density to water density |
| Red blood cells | Amount of red blood cells in urine |
| Bacteria | Identification of kidney infection, if present, and bacterial levels |
| Blood glucose random | Measurement of random blood glucose (sugar) levels |
| Blood urea | Amount of urea nitrogen in the blood |
| Serum creatinine | Presence of creatinine in blood cells |
| Sodium | Volume of sodium levels in the blood |
| Potassium | Quantity of potassium levels in the blood |
| Hemoglobin | Protein in red blood cells, with specific details |
| White blood cells count | Calculation of the number of white blood cells in the blood |
| Hypertension | Indication of high blood pressure |
| Coronary artery disease | Detects the presence of heart disease that may affect kidney function. |
| Appetite | Measures the detection of appetite loss |

**Table 3**. Performance exploration comparing the proposed feature selection algorithms

| | Accuracy | Recall | Specificity |
|---|---|---|---|
| GWO | 98.90 | 98.45 | 93.50 |
| BPSO | 95.40 | 94.35 | 93.32 |
| PSO | 93.60 | 91.45 | 90.22 |
| GA | 88.30 | 85..42 | 89.32 |

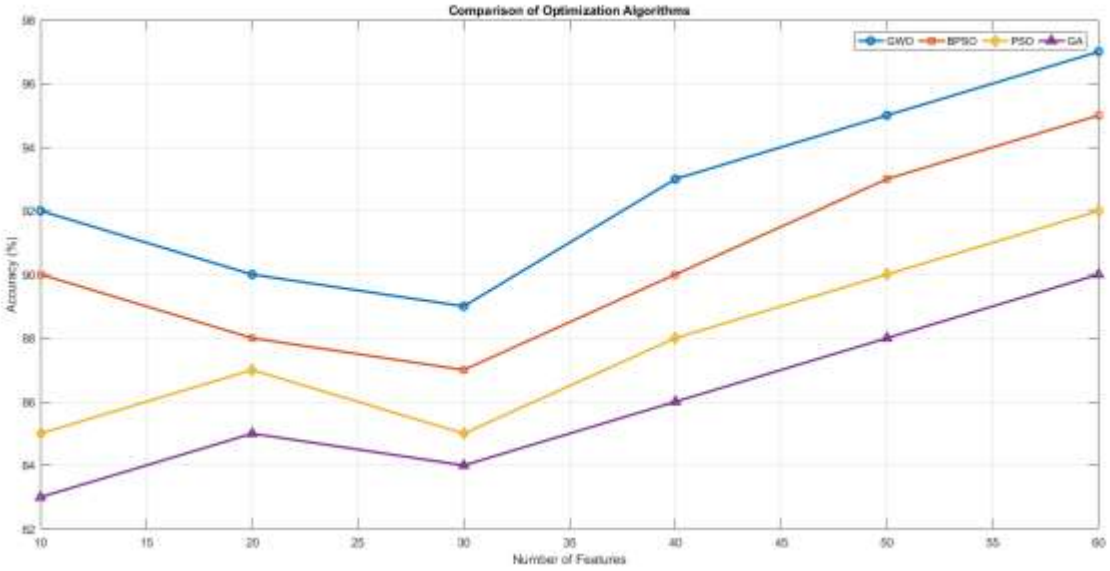


**Figure 1.** Accuracy exploration of feature selection processes

The performance comparison across different machine learning models (AdaBoost, KNN, Naive Bayes (NB), Perceptron, SVM, and DNN) shown in table 6, is measured using five key metrics: Accuracy, Kappa, Precision/Specificity, Recall, and F-score. Here's a brief summary: Accuracy Comparison: In terms of accuracy, the DNN model performs the best with an accuracy of 0.90, while SVM is the closest second one with 0.87. The perceptron has the lowest accuracy (0.78).

Kappa Comparison: Also in terms of Kappa index, the DNN model takes first place with an index of 0.88 (the level of agreement is high), the SVM is the second with an index of 0.85. Perceptron has the lowest value of Kappa (0.70).

Precision/Specificity Comparison: In terms of precision, DNN performs the best with an index of 0.92 and thus has the fewest false-positive results compared with others; SVM is the closest second with an index of 0.90. Perceptron has the lowest score in this aspect (0.75).

Recall Comparison: DNN demonstrates the best result in this case as well with an index of 0.91 (the model has the most number of true positive results). Perceptron has the lowest recall (0.70).

F-score Comparison: As a harmonic mean of precision and recall, this score is also the highest for DNN (0.90), and the lowest for Perceptron (0.73).

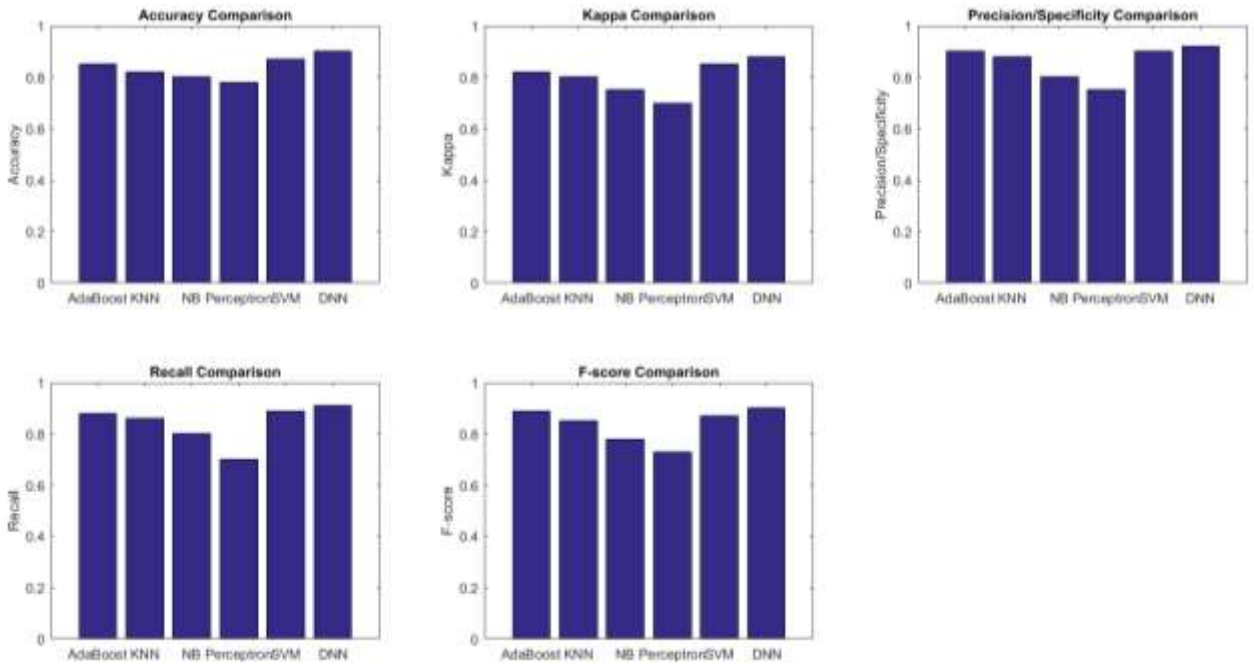


**Figure 2.** Comparative analysis of CKD medical data set classification algorithms

**Table 4.** Classification results of the proposed and existing algorithms

| | AdaBoost | KNN | NB | Perceptron | SVM | DNN |
|---|---|---|---|---|---|---|
| Accuracy | 0.9270 | .9065 | 0.8488 | 0.8304 | .9302 | .9688 |
| Kappa | 0.8460 | .07920 | 0.6738 | 0.6600 | 0.8447 | 0.6056 |
| Specificity | 0.9370 | 0.9212 | 0.8002 | 0.9765 | 0.9462 | 0.7922 |
| Recall | 0.9545 | 0.9355 | 0.9780 | 0.7540 | 0.9664 | 1.010 |
| F-score | 0.9456 | .9270 | .8808 | .08510 | 0.9434 | 0.8844 |

## 5. Discussion

The purpose of this research is to propose a hybrid algorithm that combines the power of Grey Wolf Optimization (GWO) with the Deep Belief Neural Network (DBNN) for predicting CKD. The proposed model solves a number of issues related to CKD prediction such as high dimensionality of the medical data, efficiency and classification accuracy. The GWO algorithm helped us to select the optimal features which not only reduced the high dimensionality but also the computational complexity as well because the algorithm isolated the most predictive attributes. Then, the classifier DBNN used the selected features to reach high diagnostic accuracy. Our pro-posed model showed high accuracy in predictions with 98.9% of accuracy, significantly better than any other existing model and feature selection algorithms such as PSO, BPSO, GA.

The advantage of our model compared to earlier models is that we have got high sensitivity and specificity. These two indicators are crucial for providing high accuracy of CKD predictions. The GWO feature selection technique worked perfectly because the meta-heuristic nature of the algorithm helped to identify the relevant features.

There have been earlier works done on the use of the combination of optimization algorithms with the deep learning architecture for disease detection. Our research provides an innovation in this area, where we show that the GWO-DBNN model is well suited for detecting CKD, owing to its feature selection efficiency of GWO algorithm and high-dimensional learning capability of DBNN. Although other approaches like support vector machine and simple neural networks have

proved to be effective, their computation time is higher and there is a possibility of overfitting when the number of features is large.

### 5.1 Clinical and Practical Implications

This paper concludes that the results have implications especially for early CKD diagnosis, where early and accurate diagnosis is necessary to stop the progression of the disease. As mentioned above, the sensitivity of the proposed algorithm is very high, and hence it would be able to diagnose any early signs of CKD. Therefore, the algorithm can assist medical practitioners in the early diagnosis of the disease. Additionally, the reduced computational time using GWO implies that the model can be used in real time.

### 5.2 Limitations and Future Work

However, there are a number of constraints that need to be addressed in this regard. First, the model is trained on a small dataset, which can affect its applicability to wider contexts. Further research needs to verify this approach through bigger and more diversified datasets to confirm its efficiency. Second, it would be reasonable to investigate some other deep learning models or use a different combination of optimization techniques. Finally, this predictive model can be implemented into patient-monitoring tools to assess kidneys constantly.

## 6. Conclusion

The current research suggests that chronic kidney disease analysis can be made better through using DBNNs together with the feature selection process based on the GWO approach. Thanks to the fact that our approach deals effectively with the missing data problem and focuses on important features only, our model demonstrates superiority in terms of diagnostic capabilities compared to previous approaches. The results show the efficiency of our approach in reducing the computational complexity and time required for categorizing medical data by providing higher sensitivity, specificity, kappa, F-score, and accuracy. With the help of early disease detection, these results can improve patients' treatment and contribute to development of new machine learning-based medical diagnostic models.

## Data Availability Statement

N/A

## References

[1] Elkholy, H. M., Elrashid, A., Afify, N., Moustafa, A. F., & Saleh, A. M. N. (2021). Effect of Aerobic Exercises in Prevention of Bone Mass Loss Post Bariatric Surgeries. *The Egyptian Journal of Hospital Medicine, 82*(4), 632–636.

[2] Gupta, R., Woo, K., & Jeniann, A. Y. (2021, March). Epidemiology of end-stage kidney disease. In *Seminars in Vascular Surgery* (Vol. 34, No. 1, pp. 71–78). WB Saunders.

[3] Dhaka, A. S., Rani, R., Jaiswal, G., Sharma, A., & Dev, A. (2024, November). Advancing Chronic Kidney Disease Detection: Improving Biomarkers and CT Imaging for Enhanced Detection. In *2024 International Conference on Integrated Intelligence and Communication Systems (ICIICS)* (pp. 1–7). IEEE.

[4] Dutta, S., & Bandyopadhyay, S. K. (2020). Chronic kidney disease prediction using neural approach. *medRxiv*, 2020-06.

[5] Yan, X., Li, X., Lu, Y., Ma, D., Mou, S., Cheng, Z., ... & Hu, G. (2022). Establishment and Evaluation of Artificial Intelligence-Based Prediction Models for Chronic Kidney Disease under the Background of Big Data. *Evidence-Based Complementary and Alternative Medicine, 2022*(1), 6561721.

[6] Chaitanya, S. M. K., & Rajesh Kumar, P. (2019). Oppositional gravitational search algorithm and artificial neural network-based classification of kidney images. *Journal of Intelligent Systems, 29*(1), 485–496.

[7] Boukenze, B., Haqiq, A., & Mousannif, H. (2016, May). Predicting chronic kidney failure disease using data mining techniques. In *International Symposium on Ubiquitous Networking* (pp. 701–712). Singapore: Springer Nature Singapore.

[8] Abdelaziz, E. H., Ismail, R., Mabrouk, M. S., & Amin, E. (2024). Multi-omics data integration and analysis pipeline for precision medicine: Systematic review. *Computational Biology and Chemistry, 113*, 108254.

[9] Wang, L. (2024). Mammography with deep learning for breast cancer detection. *Frontiers in Oncology, 14*, 1281922.

[10] Asif, S., Wenhui, Y., ur-Rehman, S., ul-ain, Q., Amjad, K., Yueyang, Y., ... & Awais, M. (2025). Advancements and prospects of machine learning in medical diagnostics: Unveiling the future of diagnostic precision. *Archives of Computational Methods in Engineering, 32*(2), 853–883.

[11] Fatima, A., Fatima, S., & Kiran, A. (2024). Big data and machine learning based early chronic kidney disease prediction. *Journal of Scientific Research and Technology*, 22–28.

[12] Saha, R., Saif, S., Khan, T., & Alabdultif, A. (2026). A critical systematic review of ensemble learning methods for predictive healthcare. *Discover Computing, 29*(1), 117.

[13] Kamel, S. R., & Yaghoubzadeh, R. (2021). Feature selection using grasshopper optimization algorithm in diagnosis of diabetes disease. *Informatics in Medicine Unlocked, 26*, 100707.

[14] Munguía-Siu, A., Vergara, I., & Espinoza-Rodríguez, J. H. (2024). The use of hybrid CNN-RNN deep learning models to discriminate tumor tissue in dynamic breast thermography. *Journal of Imaging, 10*(12), 329.

[15] Wang, Y., Wang, C., Luo, L., & Zhou, Z. (2019, July). Image classification based on transfer learning of convolutional neural network. In *2019 Chinese Control Conference (CCC)* (pp. 7506–7510). IEEE.

[16] Khan, A. A., Arslan, M., Tanzil, A., Bhatty, R. A., Khalid, M. A. U., & Khan, A. H. (2024). Classification of colon cancer using deep learning techniques on histopathological images. *Migration Letters, 21*(S11), 449–463.

[17] Liu, Y., As' arry, A., Hassan, M. K., Hairuddin, A. A., & Mohamad, H. (2024). Review of the grey wolf optimization algorithm: Variants and applications. *Neural Computing and Applications, 36*(6), 2713–2735.

[18] Humayun, U., Yaseen, M. T., Shahwaiz, A., & Iftikhar, A. (2024). Deep learning approaches for brain tumor detection and segmentation in MRI imaging. *Journal of Computing & Biomedical Informatics, 8*(01).