

Detection of Chronic Kidney Disease using Machine Learning Algorithms

Abstraction

Chronic Kidney Disease (CKD) is a serious health issue that requires early and accurate diagnosis for effective treatment.

This research dives into the world of machine learning, exploring how these smart algorithms can be better than traditional methods in identifying CKD.

We're looking at four different methods: Logistic Regression, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Neural Networks(mlp).

The goal is to figure out which one does the job best, considering accuracy, efficiency, and early detection.

This study is crucial because it can improve the way we diagnose and treat CKD, making a real difference in patient care.

Introduction

Chronic Kidney Disease (CKD) is a big health challenge, and the usual ways we diagnose it may not be the most accurate.

That's why we're turning to smart computer programs, using Logistic Regression, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Neural Networks to see if they can do better.

Each of these methods has its own strengths, from being easy to understand to being really good at finding patterns in data.

The reason we're doing this is simple: we want to find the best way to diagnose CKD early and accurately, so patients can get the right treatment when they need it. This research is all about making healthcare smarter and more effective.

Title : 1

ANALYSIS OF CHRONIC KIDNEY DISEASE (CKD) BY LOGISTIC REGRESSION COMPARING WITH TRADITIONAL

Problem (P):

Patients presenting with symptoms related to a specific disease (e.g., Chronic Kidney Disease).

Intervention (I):

Normal Method: Traditional diagnostic methods or expert judgment.

Comparison (C):

Logistic Regression: A machine learning algorithm used for disease diagnosis. Evaluating the effectiveness of Logistic Regression in comparison to conventional normal diagnostic methods.

Outcome (O):

Accuracy, efficiency, and reliability of disease diagnosis, including the ability to detect conditions at an early stage.

Problem & Interventions

Patients presenting with symptoms related to Chronic Kidney Disease (CKD), requiring accurate and early diagnosis for effective intervention and treatment.

- **Normal Method:** Traditional diagnostic methods or expert judgment.
- **Logistic Regression:** Implementing a machine learning algorithm specifically designed for disease diagnosis.

Evaluating the effectiveness of Logistic Regression in comparison to conventional normal diagnostic methods

Comparison

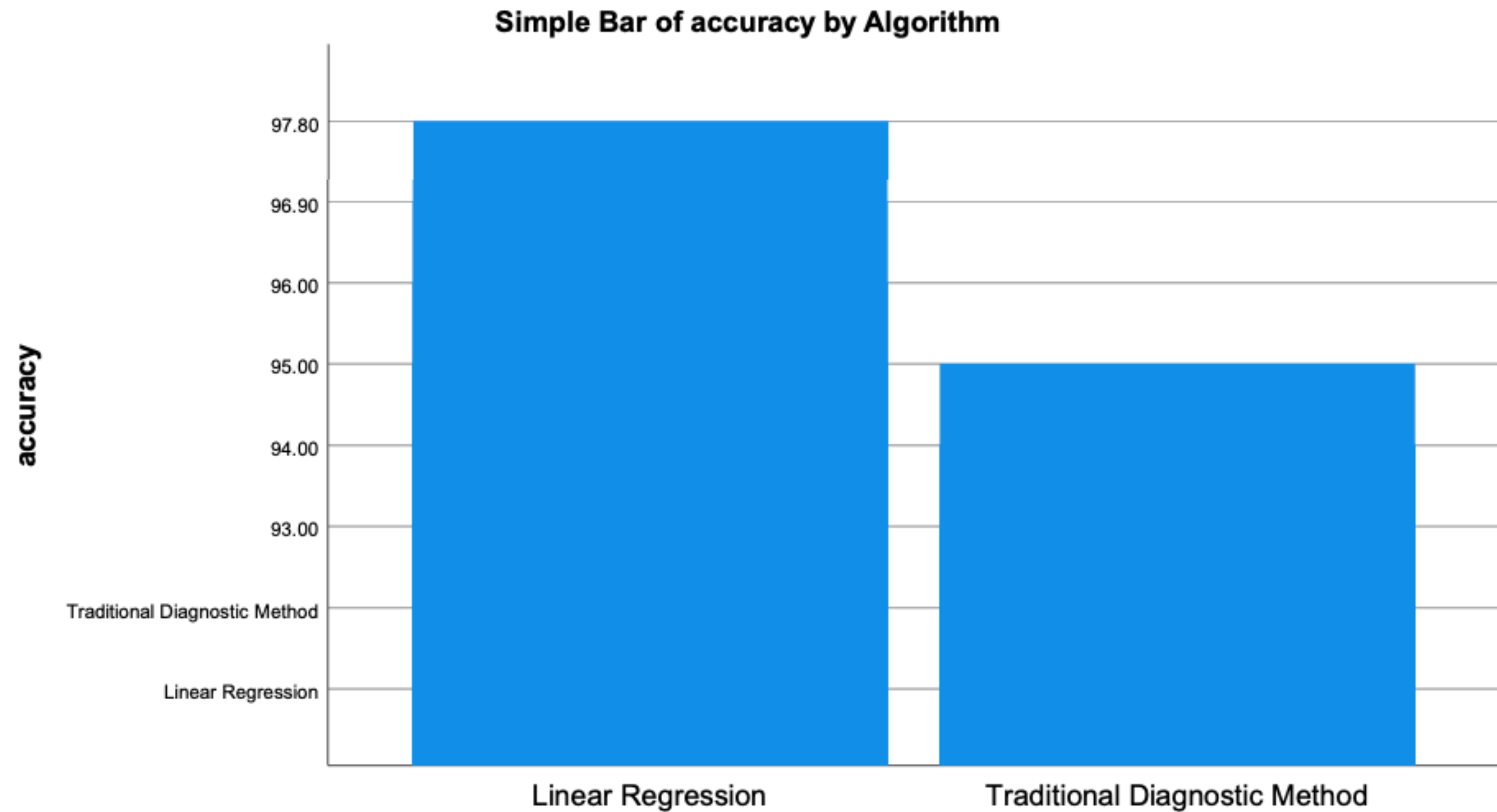
Traditional algorithm:

Conventional clinical evaluations and expert judgment are often used in traditional CKD diagnostic procedures. Based on a patient's medical history, symptoms, and the outcomes of standard laboratory testing, doctors evaluate them. These conventional methods, which frequently rely on rule-based decision-making, could not be complex enough to efficiently examine big datasets and spot minute patterns suggestive of early-stage CKD. Despite their established status and role as the foundation for medical diagnosis, their limits become evident when dealing with the growing amount of healthcare data and the requirement for more complex forecasts.

Logistic algorithm:

Conversely, logistic regression is a machine learning method that can be used for the diagnosis of chronic kidney disease (CKD) because it is well-suited for binary classification tasks. Logistic regression is used in this project to determine a linear relationship between the binary outcome variable (CKD or non-CKD) and input features (blood urea, serum creatinine, etc.). Based on a logistic function, the method generates predictions and estimates probability.

Outcome



Title 2

Problem (P):

Patients exhibiting symptoms or risk factors related to a particular disease (e.g., Chronic Kidney Disease).

Intervention (I):

Logistic Regression: A machine learning algorithm used for disease diagnosis.

Support Vector Machine (SVM): A machine learning algorithm utilized for disease diagnosis.

Comparison (C):

Assessing the effectiveness and performance of Logistic Regression in comparison to Support Vector Machine (SVM) for disease diagnosis.

Outcome (O):

Accuracy, sensitivity, specificity, and computational efficiency in disease diagnosis, particularly the ability to identify diseases at an early stage.

Problem and Interventions

Patients exhibiting symptoms or risk factors related to CKD, necessitating accurate and timely diagnosis for appropriate medical intervention.

- **Logistic Regression:** Traditional machine learning algorithm used for disease diagnosis.
- **Support Vector Machine (SVM):** Advanced machine learning algorithm designed for effective disease diagnosis.

Assessing the effectiveness and performance of Logistic Regression in comparison to Support Vector Machine (SVM) for CKD diagnosis.

Comparison

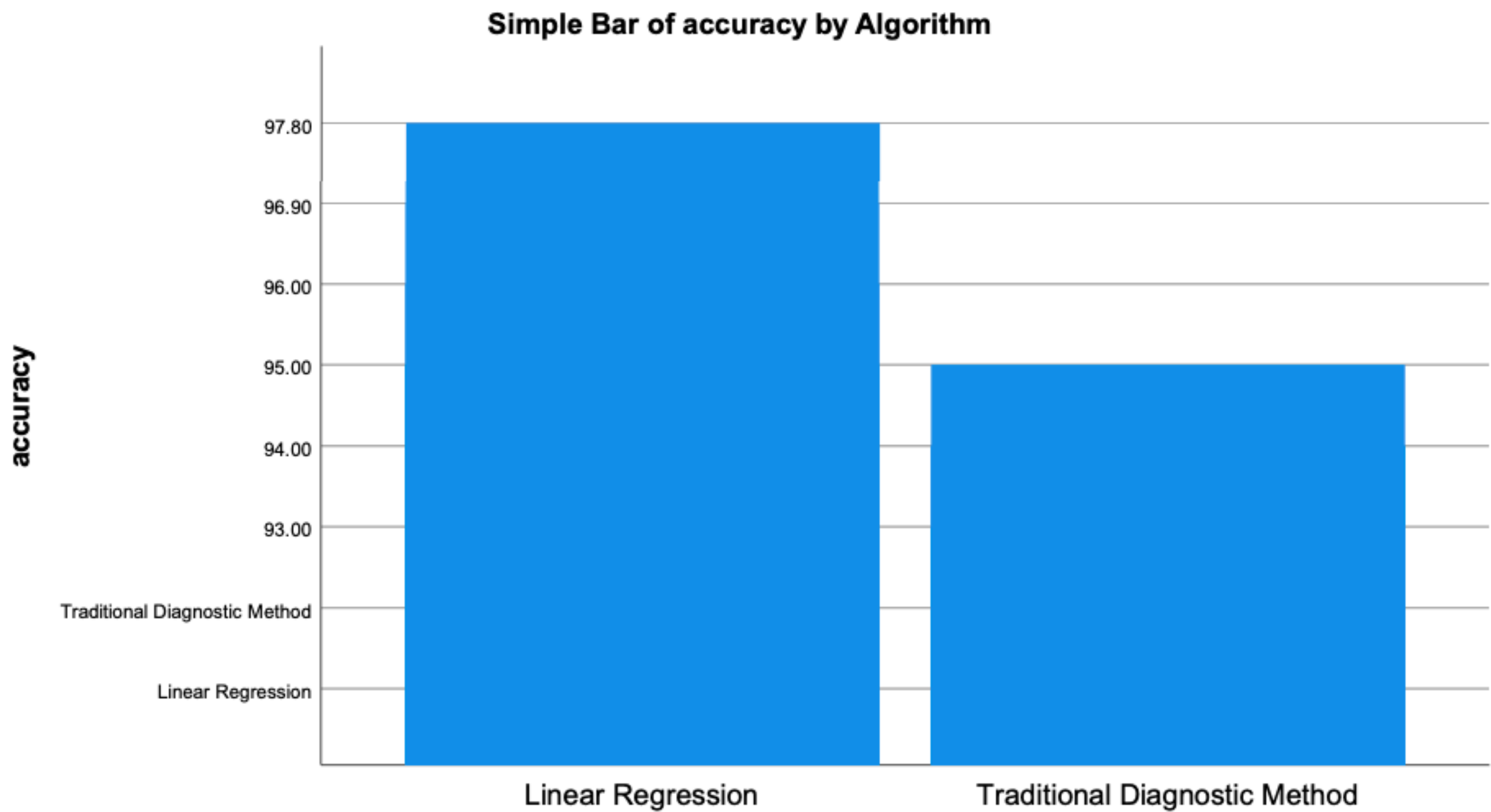
Logistic Regression Algorithm:

When applied to a dataset containing characteristics like age, blood pressure (bp), specific gravity (sg), albumin (al), sugar (su), and others, logistic regression (LR) is a useful technique for diagnosing chronic kidney disease (CKD). A confusion matrix is used to interpret the model's output, which shows true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). With a precision of 96% for non-CKD cases and 97% for CKD cases, the LR model shows excellent accuracy in the context of CKD diagnosis.

Support Vector Machine (SVM) Algorithm:

It turns out that Support Vector Machine (SVM) is a reliable and efficient diagnostic tool for Chronic Kidney Disease (CKD). As seen by the confusion matrix and classification report, the SVM model performs admirably after being trained on a dataset including features pertaining to clinical parameters. With 28 true negatives, 50 true positives, and few incorrect predictions (two false positives and zero false negatives), the matrix shows a well-balanced prediction.

Outcome



Title 3

Problem (P):

Patients exhibiting symptoms or risk factors related to a particular disease (e.g., Chronic Kidney Disease).

Intervention (I):

Support Vector Machine (SVM): A machine learning algorithm used for disease diagnosis.

K-Nearest Neighbors (KNN): A machine learning algorithm utilized for disease diagnosis.

Comparison (C):

Assessing the effectiveness and performance of Support Vector Machine (SVM) in comparison to K-Nearest Neighbors (KNN) for disease diagnosis.

Outcome (O):
Accuracy, sensitivity, specificity, and computational efficiency in disease diagnosis, particularly the ability to identify diseases at an early stage.

Problem and Interventions

Patients exhibiting symptoms or risk factors associated with CKD, requiring precise and early disease diagnosis for optimal medical intervention.

Intervention (I):

- **Support Vector Machine (SVM):** Established machine learning algorithm utilized for disease diagnosis.
- **K-Nearest Neighbors (KNN):** Machine learning algorithm known for its simplicity and effectiveness in classification tasks.

Comparison (C):

Assessing the effectiveness and performance of Support Vector Machine (SVM) in comparison to K-Nearest Neighbors (KNN) for CKD diagnosis.

Comparison

Support Vector Machine (SVM) Algorithm:

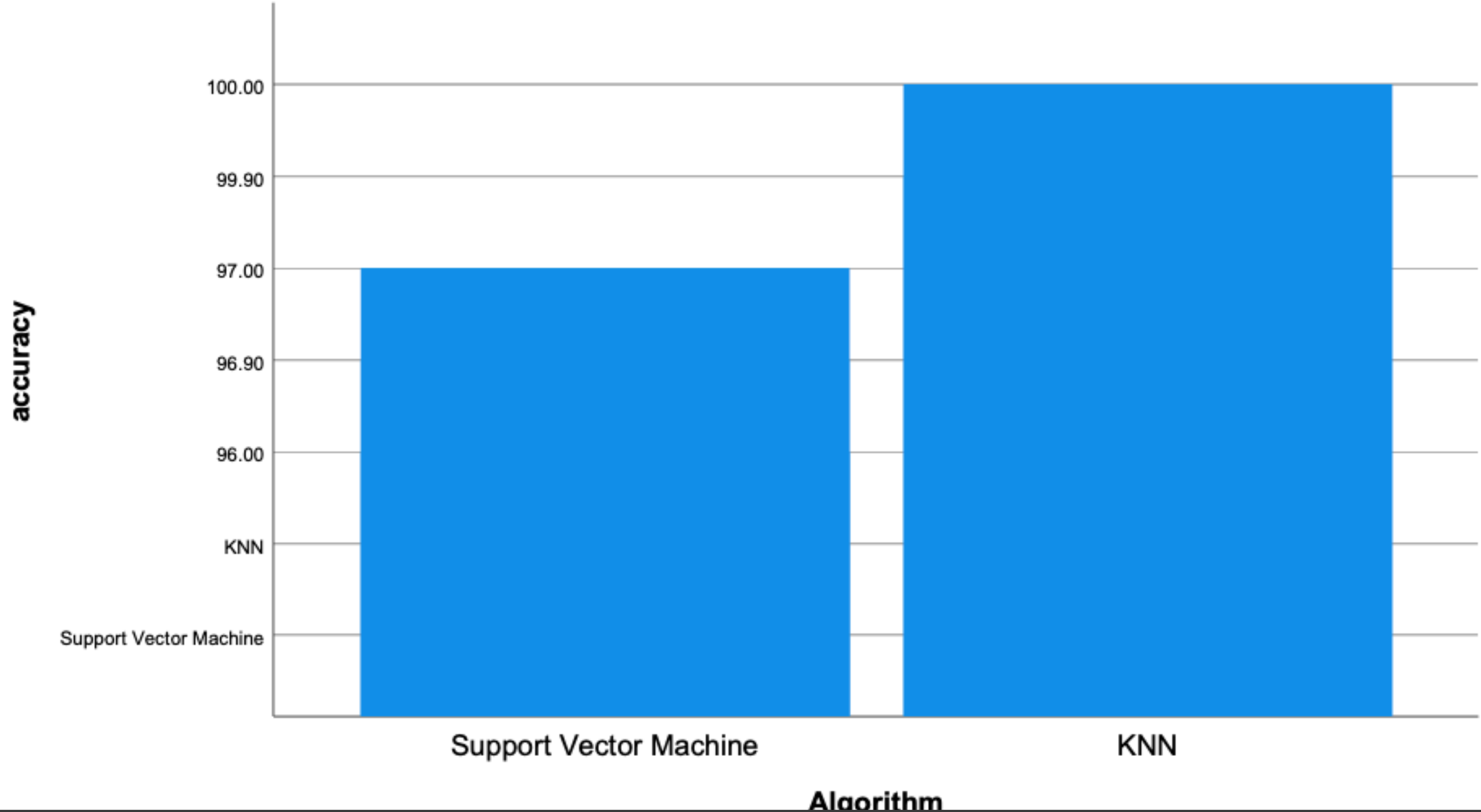
The Support Vector Machine (SVM) proved to be an effective method in the CKD diagnosis study, helping to accurately classify patients based on a variety of criteria. SVM is used here to take advantage of its capacity to identify the best hyperplanes in high-dimensional feature spaces

K-Nearest Neighbors (KNN) Algorithm:

K-Nearest Neighbors (KNN) was essential in utilizing its non-parametric, instance-based learning methodology in the CKD diagnosis investigation. The dataset served as the basis for KNN's application and included a variety of patient variables, including age, blood pressure, and biochemical markers.

Outcome

Simple Bar of accuracy by Algorithm



Title 4

Problem(P):

Patients or individuals with medical data used for disease classification or diagnosis.

Intervention (I):

K-Nearest Neighbors (KNN): A machine learning algorithm for disease classification.

Neural Networks (NN): A deep learning technique employed for disease classification.

Comparison (C):

Evaluating the performance and characteristics of K-Nearest Neighbors (KNN) in comparison to Neural Networks (NN) for disease classification tasks.

Outcome (O):

Accuracy, precision, recall, F1-score, and computational efficiency in disease classification, with a focus on their ability to correctly classify diseases based on medical data.

Problem and Interventions

Patients or individuals with medical data used for disease classification or diagnosis, specifically related to Chronic Kidney Disease (CKD).

- **K-Nearest Neighbors (KNN):** A machine learning algorithm utilized for disease classification.
- **Neural Networks (NN):** Deep learning technique employed for disease classification tasks.

Evaluating the performance and characteristics of K-Nearest Neighbors (KNN) in comparison to Neural Networks (NN) for disease classification tasks related to CKD.

Comparison

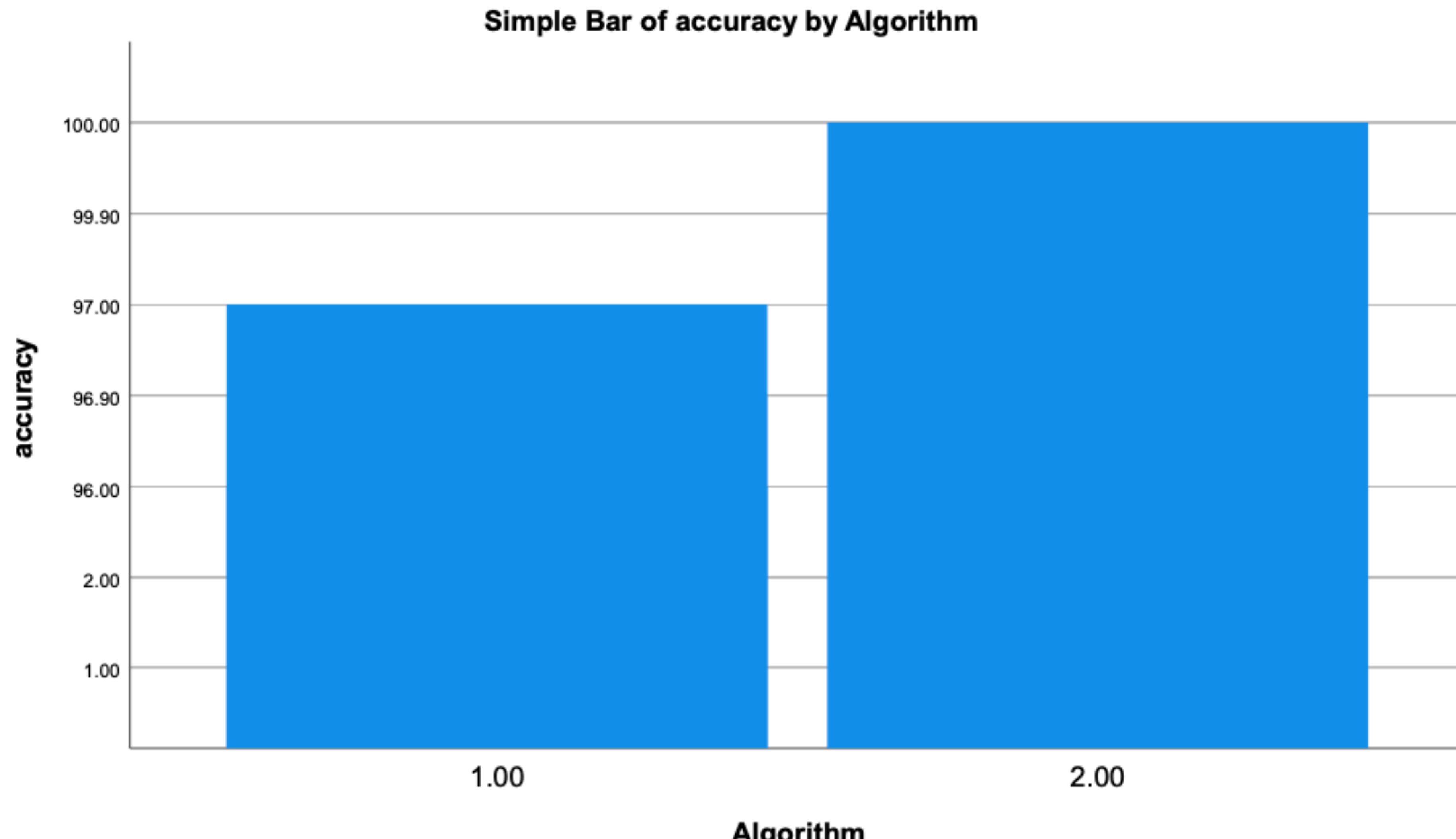
K-Nearest Neighbors (KNN) for CKD Classification:

A non-parametric, instance-based learning method called K-Nearest Neighbors is utilized for classification problems like diagnosing Chronic Kidney Disease (CKD). The closeness of data points in feature space is the foundational idea of KNN. The technique uses the presumption that cases with comparable medical characteristics are likely to be in the same class when it comes to CKD classification

Neural Networks (NN) using MLPClassifier for CKD Classification:

For CKD categorisation, neural networks—more especially, the Multi-Layer Perceptron (MLP) architecture—are an effective deep learning method. Input, hidden, and output layers are among the interconnected layers of nodes (neurons) that make up MLPClassifier. Every link has a weight attached to it, and during training the model learns to reduce the discrepancy between expected and actual results.

Outcomes



Conclusion

In the exploration of Chronic Kidney Disease (CKD) diagnosis, the application of K-Nearest Neighbors (KNN) has emerged as a notable contender, showcasing the highest accuracy among the evaluated algorithms. Through a comparative analysis with other ML algorithms KNN demonstrated competitive diagnostic performance. The findings suggest that, in the context of CKD, KNN excels in accurately identifying patients with the disease.

The high accuracy achieved by KNN signifies its proficiency in leveraging patient data to make precise diagnostic decisions.

Thank you

K.muthu-192121022