**Chronic Kieny Disease Predict**

**Problem Statement:**

The goal is to develop a machine learning model to predict Chronic Kidney Disease (CKD) based on patient medical data. The system will analyze features like age, blood pressure, glucose levels, and other clinical parameters. Early detection of CKD can help in timely intervention and better management of the disease. The solution should ensure accuracy, scalability, and ease of use for healthcare providers. Data preprocessing, feature selection, and model evaluation are critical to achieving this objective.

**Basic Dataset Info:**

Total number of Row : 399

Total number of Column: 25

**Pre-Processing:**

The code uses **StandardScaler** from sklearn.preprocessing to standardize the feature data in x\_train and x\_test. Standardization involves scaling the data such that each feature has a mean of 0 and a standard deviation of 1. The fit\_transform() method computes the necessary scaling parameters (mean and standard deviation) from the x\_train data and applies the transformation. The transform() method then applies this scaling to x\_test using the parameters learned from x\_train, ensuring consistent scaling across both datasets. This preprocessing step helps machine learning models perform better by normalizing feature magnitudes and improving convergence.

**Final Model:**

**Algorithm:** KNeighborsClassifier

**Classification Report:**

**Accuracy: 0.96**

A screenshot of a computer screen

Description automatically generated

**Precision**: Proportion of correctly predicted positives out of all predicted positives (e.g., for True, 94% of predictions were accurate).

**Recall**: Proportion of correctly predicted positives out of all actual positives (e.g., True predictions correctly covered 100% of actual positives).

**F1-score**: Harmonic mean of precision and recall, balancing both metrics (e.g., True has an F1-score of 0.97).

**Support**: The number of actual occurrences of each class (e.g., 56 for False and 77 for True).

**Accuracy**: Overall correctness of the model, with 96% of predictions accurate across all classes

**Justify The Final Model:**

The dataset consists of numerical input features, making it suitable for machine learning algorithms.

The output is categorical, indicating that a classification algorithm is appropriate for this problem.

The selected model, KNeighborsClassifier, achieves an overall accuracy of 96%, demonstrating strong performance in classifying the given data.

The high precision, recall, and F1-scores for both classes further validate the model’s reliability.

Based on the accuracy and metrics, KNeighborsClassifier effectively addresses the classification task.