**PREDICTING CORONARY HEART DISEASE USING THE KNN ALGORITHM**

**BY**

**ABISOYE D. DAMILOLA**

**Introduction**

This report documents the development of a predictive model from a dataset using KNN algorithm to determine the likelihood of an individual developing coronary heart disease within a 10-year period.

**Data Preprocessing**

The dataset underwent extensive cleaning and Exploratory Data Analysis (EDA), including:

* Handling missing values through filling and stripping text off a column
* Casting and replacing text values with numerical values for machine learning compatibility.

**Data Split**

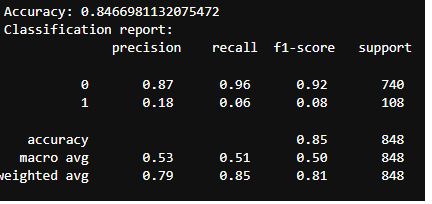
The dataset was divided into training (80%) and testing sets (20%) using the train\_test\_split function from the sklearn library’

**Model Training**

A K-Nearest Neighbors (KNN) Classifier was trained on the dataset. This algorithm selects the closest values in the dataset to predict an outcome.

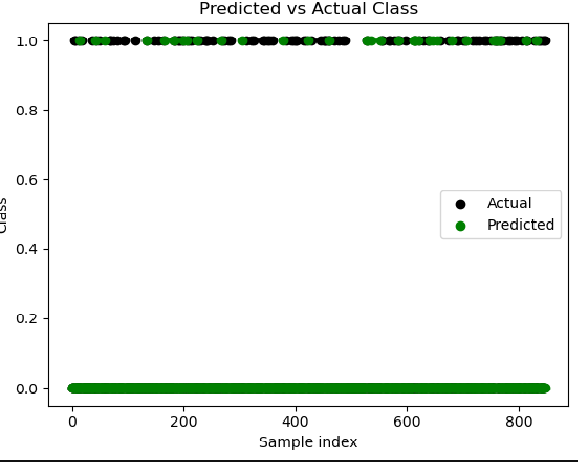
**Model Evaluation**

The model achieved an accuracy of 85% (Fig 1)



**Testing and Prediction**

The model was tested against the test dataset earlier divided, yielding a predicted answer of “0”, indicating a low likelihood of developing coronary heart disease in 10 years. (Fig 2)

****

**Conclusion**

The trained KNN model demonstrates a high accuracy in predicting coronary heart disease risk, providing a valuable tool for healthcare professionals to identify at-risk individuals and implement preventive measures.