Oruru Ndu-Nwogu T00686733

Project Name: Medical Diagnosis Using Ensemble Methods.

IDE: RStudio.

Languages:

R: To be used for the analysis, visualization and preparation of the data.
Python: To be used to create and compare the machine learning models.

Overview: Develop a medical diagnosis system that combines multiple machine learning algorithms including **KNN, SVM, Random Forest, and Decision Trees.**

Use patient data such as symptoms, medical history, and test results to predict the likelihood of various diseases or conditions.

If possible, add a deep learning model towards the end and compare it to see if it's better.

Goal: Evaluate the ensemble's performance in terms of **accuracy, precision, sensitivity, and specificity.**

Ensemble Method to be used:
Stacking: involves training a new model to combine the predictions of several base models. The base models are trained on the full dataset and then the meta-model is trained on the outputs of the base models as features.
Applicability: Effective if the base models are significantly different. This technique can leverage the strength of each base model to improve overall performance.

Dataset Type:
Single dataset: focuses on a specific disease

Dataset: Heart Disease

Dataset Link(from Kaggle):
https://www.kaggle.com/datasets/fedesoriano/heart-failure-prediction

Publication: R-PUB