Heart Risk Classification using Logistic Regression and KNN

Problem Statement

Cardiovascular diseases are a leading cause of mortality worldwide. Predicting the risk of heart problems can be crucial for early intervention and preventive measures. In this assignment you are tasked with developing a predictive model using computational intelligence techniques to assess the likelihood of an individual experiencing heart problems.

Instructions

1) Data Collection:
   - Access the data from any one of the following provided sources:
     - ScienceDirect Heart Disease Prediction Dataset (Table 1) (https://www.sciencedirect.com/science/article/pii/S2352914819301996).
     - Kaggle Heart Attack Prediction Dataset (https://www.kaggle.com/datasets/iamsouravbanerjee/heart-attack-prediction-dataset).
   - Explore the dataset to understand the features and their distributions.
   - Normalize, standardize, preprocess the data as needed.
   - Identify which features are the most significant and least significant. Justify your answer

2) By using Logistic Regression and KNN, develop models to predict heart disease risk. Compare the performance of both models. Is one of the models more superior to the other? Justify the your answer.