CS 513 B – KDD PROJECT PROPOSAL: (Heart Attack Analysis)

CS 513 B - KDD Project Proposal: Heart Attack Analysis

Project Group No: 5

Problem Statement:

Heart Attack Analysis is a problem that involves predicting the likelihood of a person experiencing a heart attack based on various risk factors. The problem statement typically involves using a dataset of patient information, such as age, gender, cholesterol levels, blood pressure, and other health metrics, to develop a predictive model that can accurately identify individuals at high risk of a heart attack. The goal of this analysis is to enable healthcare professionals to identify patients who may require additional screening or treatment to prevent a potentially life-threatening heart attack.

Dataset: (Description)

The dataset comprises 21 features in the form of columns, out of which we may opt to use the essential features only, during implementation.

Source of Dataset:

https://www.kaggle.com/datasets/alexteboul/heart-disease-health-indicators-dataset?resource=download

Implementation Strategy and algorithms used: (List different models)

We have decided to implement and compare 8 other models among four different group members. We have chosen a few models from our course and a few from outside the course. The following are the models selected by us:

1. K-nearest neighbor
2. AdaBoost Classifier
3. Random Forest
4. Support Vector Machine
5. Gaussian Naive Bayes
6. Logistic Regression
7. Decision tree
8. Artificial Neural network

Model metrics and Evaluation: (List evaluation metrics like AUC-ROC, Confusion matrix, F1, Recall, and Precision)

Evaluation of different models used in the project are

- Confusion Matrix
- F1

- Recall
- Precision

Team Members: Group 1

1. Member 1 Aman Gupta -20018346
2. Member 2 Parth Rajeshkumar Patel - 20010839
3. Member 3 Ashwin Dhanasamy
4. Member 4 Buruju Sowmya 20016102