MLProject

Predicting and Classifying Heart Disease Subtypes Using Machine Learning: Early Detection of SCA, MI, Heart Failure, and CAD

Overview

This repository contains a detailed project on predicting and classifying heart disease subtypes using machine learning. The project explores various supervised and unsupervised models, evaluates their performance, and provides insights into the prediction of heart disease events. The focus is on applying machine learning algorithms effectively to classify heart disease conditions, leveraging clinical data.

Roadmap

Project Proposal:

- Introduction to the problem statement.
- Objectives and scope of the project.

Midterm Checkpoint:

- Progress report with early insights and challenges faced.
- Initial findings from the exploratory data analysis (EDA).

Final Report:

- Comprehensive documentation of the project's methodology, results, and conclusions.
- Includes model performance evaluation and future work.

Model Implementations:

Three implemented models available as Jupyter notebooks:

1. **KMeans Clustering**: An unsupervised learning approach for exploratory data analysis.

2. **Logistic Regression**: A supervised classification model with emphasis on interpretability.

3. **Support Vector Machine (SVM)**: A robust supervised classification method for handling complex data.

Contents

Jupyter Notebooks

- [KMEANS_CLUSTERING.ipynb](): Implementation of K-Means for exploratory analysis and clustering.
- [LOGISTIC_REGRESSION.ipynb](): Logistic regression model for binary and multiclass classification.
- [SVM.ipynb](): SVM implementation for predictive analysis and classification.

Reports

- [Project Proposal]()
- [Midterm Checkpoint]()
- [Final Report]()

Supporting Files

- **Gantt Chart**: Timeline and milestones of the project ([view here]()).
- **README.md**: This file, detailing the structure of the repository and instructions.

Authors

- Maria Jose Jimenez
- Sofia Fraija
- David Silvera
- Carlota Huertas
- Kushaal Raju Palasamudrum