Data Science

Project Proposal

Member#1: Tayyab Shafiq **Roll#** cs201153

Member#2: Taha Khalfe **Roll#** cs201096

Project Title

Predicting the Survival of Patients with Heart Failure.

Introduction

Problem Statement: The project aims to address the challenge of predicting the survival of patients with heart failure using machine learning techniques. Specifically, the researchers seek to utilize available electronic medical records to quantify symptoms, body features, and clinical laboratory test values in order to predict patients' survival. By doing so, they aim to highlight patterns and correlations that may not be readily detectable by medical doctors.

Objectives:

1- To use machine learning techniques to predict the survival of patients with heart failure using available electronic medical records.
2- Identify the main features that are predictive of patient survival.

Scope and Limitations:

1- The limitations of the project include the small size of the dataset, which may limit the generalizability of the findings.
2- The dataset does not include information about the patients' physical features such as height, weight, and body mass index, which could be useful in identifying additional risk factors for cardiovascular health diseases.

Project Overview

The project aims to use machine learning techniques to predict the survival of patients with heart failure using available electronic medical records. By analyzing a dataset of medical records of 299 patients with heart failure and by identifying the specific clinical and laboratory test values that are most strongly associated with patient survival,

Data Sources

Data Types: The project contains 12 features in which 5 features are integer data type and discrete in nature, 1 is in float data type, 5 are categorical and the output column is also categorical.

Collection Methods:

Research is published on “BMC Medical Informatics and Decision Making”

https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-020-1023-5

Dataset is made publicly at “Public Library of Science research repository”

https://plos.figshare.com/articles/dataset/Survival_analysis_of_heart_failure_patients_A_case_study/5227684/1

Methodology

Machine Learning Algorithms Used:

Logistic regression with sigmoid activation.