Examining the main factors and their extent of influence to Cardiovascular diseases death events

Jingwen Liang jl4436
Peilin Li pl593
Guohan Gao gg479

Problem Statement

This project seeks to evaluate the primary factors and their impact on mortality events associated with cardiovascular disease. It will explore which specific factors contribute most significantly to the incidence of fatal outcomes in heart failure patients. This project will address two fundamental questions:

1. What are the main factors that cause Cardiovascular diseases death events?
2. How do different features influence Cardiovascular diseases death events?

By using different machine learning algorithms, this project will predict likelihood of mortality events among different features with the heart failure dataset. This provides insight into forecasting the risk of death events, helping to enhance patient care and reduce healthcare costs.

Dataset

This project intends to use a dataset comprising health records from 299 patients with heart failure, which was collected in 2015 by the Faisalabad Institute of Cardiology and the Allied Hospital. The dataset includes 13 clinical attributes per patient, categorized into three data types: binary (anaemia, diabetes, high blood pressure, sex, smoking, and death event), continuous (creatinine phosphokinase, ejection fraction, platelets, serum creatinine, and serum sodium), and integer (age and follow-up time). The 'death event' will be taken as output, indicating survival (1) or death (0) during the follow-up period, while the remaining 12 features serve as input. Out of the 299 individuals, 203 survived and 96 passed away (*Heart Failure Prediction*, n.d.). This dataset is clean with no missing entries, including the main possible factors of CVDS, which provides a comprehensive numerical foundation for analyzing the factors contributing to mortality in heart failure patients.

Significance

Cardiovascular diseases (CVDs) are the foremost cause of death globally, responsible for an estimated 17.9 million fatalities each year, accounting for 31% of all deaths worldwide. Among these, heart failure stands out as a common outcome of CVDs. The data we focus on includes 12 critical features that are pivotal in forecasting the risk of mortality due to heart failure. Addressing these can significantly impact global health outcomes and provide key insights into prevention and management strategies.

To combat the prevalence of CVDs, it is imperative to target modifiable behavioral risk factors such as tobacco usage, poor dietary choices, obesity, sedentary lifestyle, and excessive alcohol consumption. Public health measures aimed at these risk factors can effectively reduce the incidence of heart diseases. Furthermore, for those already affected by CVDs or at elevated risk due to conditions like hypertension, diabetes, or hyperlipidemia, early detection and proactive management are essential. In this realm, machine learning models hold immense promise, offering advanced tools for timely intervention and improved patient care, ultimately aiding in the reduction of CVD-related mortality.

Early detection and intervention for CVDs can dramatically reduce long-term healthcare costs and minimize productivity loss due to illness, thereby positively impacting the economy. By investing in technologies such as machine learning models for CVD prevention and management, enterprises can enhance their technological image and competitive edge. They also stand to benefit directly through reduced healthcare insurance payouts and improved employee health levels. Consequently, our studies are not only conducive to advancing public health objectives but also align with the commercial interests of businesses, providing a compelling case for their support in the long term.

References

Heart Failure Prediction. (n.d.). Kaggle. Retrieved March 16, 2024, from

https://www.kaggle.com/datasets/andrewmvd/heart-failure-clinical-data