

## Cardiovascular Disease Data Analysis Using Machine Learning Models

**Authors :** Ranveet Saggu, Saad Bin Ahmed

**Abstract :** Cardiovascular Disease (CVD) is the leading cause of death worldwide. One of its main manifestations, myocardial infarction (commonly known as a heart attack), occurs about 750,000 times a year, caused by insufficient blood flow to a portion of the heart muscle. A quick and accurate diagnosis of a heart attack or heart failure is crucial in the treatment of the patient. The aim of this research project is to improve the prediction of cardiovascular diseases by automating risk assessment using binary classifiers. The methodology includes Exploratory Data Analysis (EDA), which helps to obtain information about the dataset with the help of visualizations and metrics. Additionally, Feature Engineering techniques is employed to address missing values, outliers, feature extraction, and normalizing the dataset. Subsequently, various classification machine learning algorithms are trained, and their accuracy along with other metrics are evaluated to identify the most efficient model in terms of processing time and predictive performance.

**Keywords :** cardiovascular disease, machine learning, decision trees, logistic regression, k-nearest neighbor, xgboost, random forest, gradient boosting

**Conference Title :** ICBS 2026 : International Conference on Biomedical Sciences

**Conference Location :** Rome, Italy

**Conference Dates :** January 18-19, 2026