HeartSense: Leveraging Machine Learning to Predict Cardiovascular Risk

Justin Lee
InspiritAI, Lexington, SC 29072 USA
justinleethirtythree@gmail.com

January 19, 2025

Outline

Abstract

Heart disease, also known as cardiovascular disease, remains one of the leading causes of mortality worldwide. Accurate prediction models are critical for advancing early detection and preventive measures. This research introduces a specialized machine learning framework to predict the severity of heart disease by using the "UCI Heart Disease Data" from Kaggle—a multivariate dataset derived from the Cleveland database.

This dataset encompasses 14 predictive attributes, including clinical and demographic factors, which were used to train and evaluate various supervised learning algorithms. Notable models included logistic regression, decision trees, random forests, and gradient-boosting machines. The highest-performing model (XGBoost) achieved an accuracy of *****PERCENTAGE*****.

This study demonstrates how machine learning can uncover nuanced patterns within medical datasets, offering actionable insights into cardiovascular health and aiding in clinical decision-making.

Introduction

- **Global Impact of Cardiovascular Diseases (CVDs):**
  - Leading cause of death worldwide (17.9 million deaths annually).
  - High economic burden on health systems globally.
- **Role of Emerging Technologies:**
  - Artificial Intelligence (AI) and Machine Learning (ML) as transformative tools.
  - Paradigm shift: reactive treatment $\rightarrow$ proactive prevention.
- **ML Advantages:**
  - Scalability and real-time predictions.
  - Analysis of diverse data sources (EHRs, genetics, lifestyle factors).
  - Improved accuracy, early detection, and personalized treatment plans.

Methodology

Dataset Overview:

- UCI Heart Disease Data from Kaggle.
- 14 predictive attributes (e.g., age, sex, cholesterol levels).
- Multivariate dataset used to evaluate multiple ML algorithms.

Supervised Learning Algorithms:

- Logistic Regression.
- Decision Trees and Random Forests.
- Gradient-Boosting Machines (e.g., XGBoost).

Model Development:

- Hyperparameter tuning for optimization.
- Cross-validation techniques for robust evaluation.

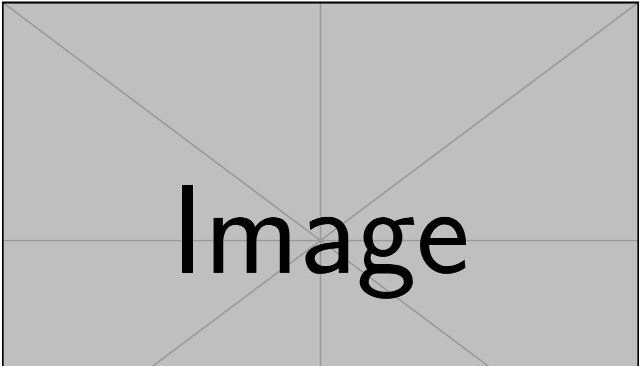
Evaluation Metrics:

- Precision, Recall, and F1 Score.
- Sensitivity and specificity to balance predictions.

Results

Key Findings

- **Best Model:** XGBoost with accuracy of ***PERCENTAGE***.
- Precision and recall metrics indicated strong model performance.
- Hyperparameter tuning significantly improved results.



Conclusion

- Machine learning offers scalable, accurate methods for heart disease prediction.
- Advanced feature engineering and systematic tuning are crucial.
- Integration into clinical workflows can enhance decision-making and patient outcomes.
- **Future Work:**
  - Explore deep learning models for more complex patterns.
  - Use larger, real-world datasets for validation.
  - Investigate model interpretability for clinical adoption.

Thank You! Questions?