



INFORMATICS  
INSTITUTE OF  
TECHNOLOGY

INFORMATICS INSTITUTE OF TECHNOLOGY

In Collaboration with

UNIVERSITY OF WESTMINSTER

**CardiacPredict: Predicting Cardiovascular Disease Risk Using Ensemble Models  
and Explainable AI**

A dissertation by

Mr. Dineth Panagoda

W1873506 / 20211146

Supervised by

Mr. Manul Singhe

Submitted in partial fulfilment of the requirements for the BEng (Hons) Software  
Engineering degree at the University of Westminster.

**April 2025**

## Abstract

Cardiovascular disease (CVD) continues to be the primary cause of global morbidity and mortality, highlighting the critical need for accurate risk prediction tools. Traditional risk models, such as the Framingham Risk Score, rely on cross-sectional data and conventional statistical techniques that often fail to capture complex interactions among risk factors. In this project, we present an ensemble machine learning framework designed to predict coronary heart disease (CHD) risk using the UCI Heart Disease dataset. Our approach incorporates advanced feature engineering techniques—such as generating interaction terms (e.g., age multiplied by systolic blood pressure), calculating differences (e.g., the gap between systolic and diastolic blood pressure), and applying logarithmic transformations—to better model non-linear relationships among variables. Individual models, including Random Forest, XGBoost, Gradient Boosting, and a Neural Network, are meticulously tuned via cross-validation and combined using stacking ensemble methods. Techniques like SMOTE are employed to address class imbalance, while explainable AI methods, particularly SHAP, provide both visual and textual insights into the contributions of key features. Performance is evaluated through ROC AUC, accuracy, precision, recall, and F1 score. This work aims to deliver a robust and interpretable CHD risk prediction tool that can enhance clinical decision-making and contribute to advancements in cardiovascular risk assessment.

### Subject Descriptors:

- Computing methodologies → Artificial intelligence → Machine learning → Ensemble methods
- Computing methodologies → Artificial intelligence → Machine learning → Explainable AI
- Applied computing → Life and medical sciences → Health informatics → Clinical decision support systems
- **Keywords:** CHD detection, machine learning, ensemble processing, Explainable AI