

DEEP VEIN THROMBOSIS

MACHINE LEARNING APPROACH

INTRODUCTION :

This project aims to leverage machine learning to improve diagnostic precision for Deep Vein Thrombosis (DVT) using data from the UK Biobank. By applying diverse machine learning techniques.



OBJECTIVE :

The goal of this project is to use machine learning to find important factors causing DVT in patients using data from the UK Biobank and improve chances of early detection.

DATABASE :

The UK Biobank is a large-scale biomedical database that is held in the UK. The database contains detailed medical information about more than 500,000 individuals.

- 500,000+ INDIVIDUALS IN THE DATASET
- 4 DIFFERENT TIME POINTS IN THE DATASET
- SPARSE DATA : MANY OF THE FEATURES ARE MISSING OR IRRELEVANT
- MIXED DATA TYPES : CATEGORICAL + CONTINUOUS DATA

MACHINE LEARNING - UTILIZED ALGORITHMS :

1. XGBoost:

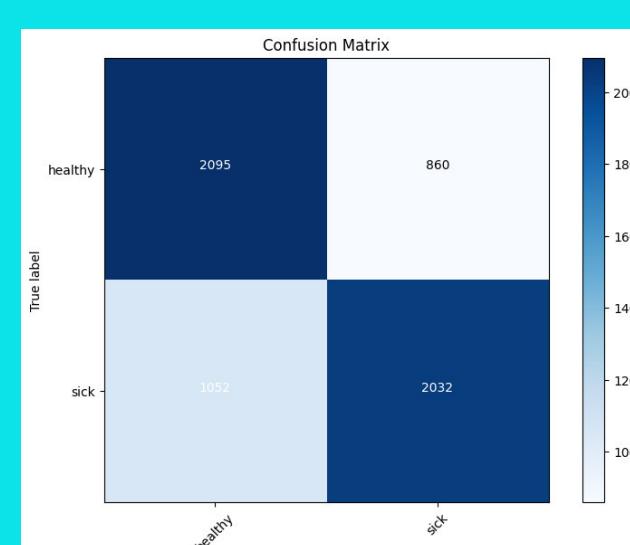
An advanced gradient boosting algorithm for powerful predictive modeling, particularly effective in handling complex datasets.

2. Logistic Regression:

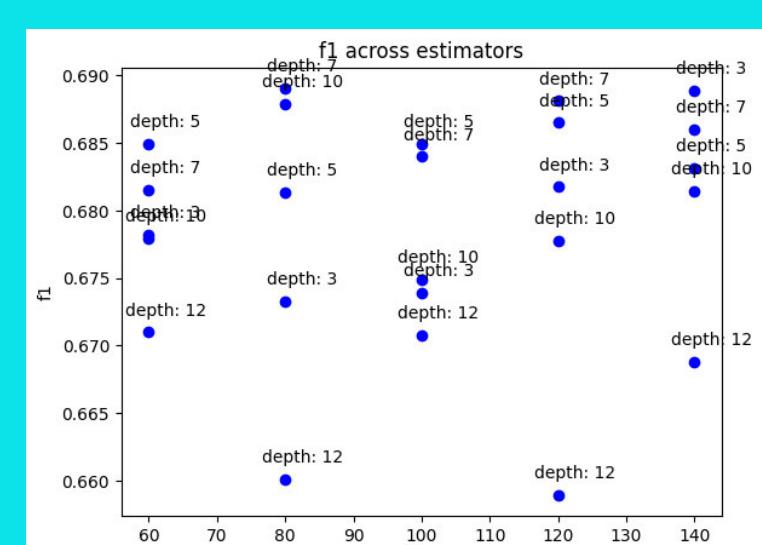
A fundamental statistical technique for binary classification, providing interpretable results and a baseline for comparison.

ANALYSIS :

Confusion Matrix



F1 Score



- CONFUSION MATRIX: VISUALIZES MODEL PERFORMANCE FOR ERROR ANALYSIS.
- F1 SCORE: BALANCES PRECISION AND RECALL VITAL FOR IMBALANCED MEDICAL DATA.

CONCLUSION :

We've successfully generated meaningful predictions, paving the way to identify factors influencing DVT risk. Our next steps involve using these insights to develop effective strategies for DVT prevention and management.

