**Ensemble Techniques for Coronary Artery Disease Risk Prediction Adapting Electronic Health Record**

A. Fatima Zohra1 , B. Dr. Onkara Perumal. P1

1 Department of Biotechnology, NIT Warangal

[FZBT21110@STUDENT.NITW.AC.IN](mailto:FZBT21110@STUDENT.NITW.AC.IN)

**ABSTRACT**

Coronary Artery Disease (CAD) is a major cardiovascular disease affecting a substantial human population globally. The morbidity, mortality, and socio-economic importance of this disease prompts for an accurate besides timely prognosis of disease risk for cost-effective management of CAD. Detection of hidden patterns in large electronic health record (EHR) datasets with a multitude of variables adopting machine learning approaches is proposed as a viable strategy.

Our aim is to develop a robust machine learning model for predicting CAD risk among individuals by employing available EHR data for cost-effective management of CAD.

In our present work we have employed an innovative embodied machine learning algorithm viz. XGBoost, Ada Boost, Gradient Boosting for enhanced prediction of CAD risk adopting EHR data. From our observation our embodied model displayed improved efficiency in terms of better prediction accuracy and consistency as corelated with contemporary machine learning approaches. Two IDEs, jupyter and Spyder were adopted for generic and modular approaches respectively. Likewise, implementation of SHAP and LIME approaches are characteristic of our present work.

Individually the models namely XGBoost, Ada Boost, Gradient Boosting have shown the accuracy of 93.57%,83.13%,86.26% respectively but after putting them into soft voting classifier the final accuracy is 92.52%.

On successful completion of the task, the developed model would be deployed on an open cloud server HEROKU for public use.

**KEY WORDS:** CAD, XGBoost, Ada Boost, Gradient Boosting, machine learning