

Improving Coronary Artery Disease Prediction: The Role of GAN-Simulated Genetic Variants in Machine Learning Models

Publisher: IEEE

Cite This

PDF

Fatima Ansari ; Ayushi Singh ; Rahul

All Authors

26 Full Text Views

Abstract

Document Sections

I. Introduction

II. Current Methods for Cad Risk Prediction

III. Gan in Healthcare and Genomics

IV. Limitations of Gan-Generated Genetic Variants

V. Future Prospects

Show Full Outline

Authors

Figures

References

Keywords

Metrics

More Like This

Abstract:

Coronary artery disease (CAD) is the leading cause of global mortality through the narrowing of coronary arteries from plaque buildup, which restricts blood flow to the heart. Recent studies indicate the genetic underpinnings of CAD and how complex genetic variations affect its occurrence. However, the scarcity and heterogeneity of genetic datasets make it difficult to predict CAD with a high degree of accuracy, particularly for rare critical variants. The combination of patient medical history, clinical variables, and genetic information through machine learning algorithms enhances the prediction of CAD. However, techniques such as Random Forest, XGBoost, and LightGBM have the potential but are limited in their predictive capabilities due to data quality and representation issues. This work proposes the use of GANs as a solution to overcome these inherent difficulties - generating synthetic genetic variants that can capture the nuances in real - world data. The GAN-simulated data closely mimics true patterns of genetic variants and hence reveals complex genetic interactions. Addition of the GAN-simulated data to the CAD prediction model significantly improves the generalization capability across different populations, reducing bias and allowing an improved estimation of risk. In conjunction with explainable AI techniques, this GAN-based approach would enable prevention for individuals at higher risk of developing CAD. Though GANs offer a lot of promise, overcoming challenges like overfitting and mode collapse as well as ethics associated with using synthetic data calls for a judicious approach for their proper application in the clinic.

Published in: 2025 International Conference on Cognitive Computing in Engineering, Communications, Sciences and Biomedical Health Informatics (IC3ECSBHI)

Date of Conference: 16-18 January 2025

DOI: 10.1109/IC3ECSBHI63591.2025.10990640

Date Added to IEEE Xplore: 20 May 2025

Publisher: IEEE

ISBN Information:

Conference Location: Greater Noida, India

I. Introduction

Coronary artery disease (CAD) is one of the most common diseases in the world, caused by the blockage of coronary arteries in the heart that reduces blood flow to the heart. Without timely medical interventions, CAD is still a major concern in the world.

Sign in to Continue Reading

Authors

Figures

References

Keywords

Metrics

PDF

Help

https://ieeexplore.ieee.org/abstract/document/10990640

1/2