

MACHINE LEARNING TO ASSESS CARDIOVASCULAR RISK IN T2DM PATIENTS

OBJECTIVES:

The aim of this analysis was to streamline cardiovascular (CV) risk assessment by reducing the number of inputs required, thereby enhancing efficiency and feasibility in clinical settings and improving physician acceptance

METHODOLOGY:



Data and Algorithm

Utilized data from the MARK study on asymptomatic T2DM patients in India and applied the XGBoost algorithm, a decision tree-based ensemble method



Dataset Processing

Divided the data into training (80%) and test (20%) sets, using SMOTE + ENN for balancing. Dimension reduction techniques identified key clinical parameters



Model Evaluation

Assessed model accuracy with a confusion matrix and explained predictions using SHAP values

KEY FINDINGS:

Dimension Reduction

XGBoost reduced parameters from 22 to 10, achieving 95.12% accuracy in predicting 10-year cardiovascular risk

Model Accuracy

The model correctly classified 95.12% of patients, with minimal misclassification

SHAP Values

SHAP values clarified each feature's impact on the model's predictions

Clinical Relevance

The reduced-parameter model simplifies cardiovascular risk prediction in Type 2 diabetes patients for clinical use