20. **MULTILAYER PERCEPTRON ARTIFICIAL NEURAL NETWORKS AND TREE MODELS AS MULTIFACTORIAL BINARY PREDICTORS OF HEART DISEASE AND FAILURE**

Jehad Amer Yasin[1].

[1] Second-year Medical Student. The University of Jordan, Amman, Jordan

https://www.youtube.com/watch?v=vlsNiqV1-28&t=5668s

BACKGROUND: Cardiovascular diseases (CVDs) are a significant global health concern, causing an estimated 17.9 million deaths annually, which represents 31% of worldwide deaths. A significant proportion of CVD deaths are due to heart attacks and strokes, with one-third of these deaths occurring prematurely in individuals under 70 years old. Heart failure is a notable event within CVDs and emerges when the heart cannot efficiently pump blood to fulfill the body's requirements. This complex syndrome's origins are multifactorial and often arise from conditions such as hypertension, diabetes, and hyperlipidemia. Large datasets with multiple features offer an opportunity for machine learning to aid in the early detection and prediction of heart failure. **METHODS**: The study employed an unmatched case-control retrospective design. Supervised machine learning models were utilized, notably Multilayer Perceptron Artificial Neural Networks (MLP-ANNs) and decision tree-based models, to predict heart failure disease using data from 918 patients. The open licensed dataset, a combination from five independent heart datasets, comprises 11 demographic and clinical features related to patient status. The MLP-ANN, equipped with a hidden layer and a hyperbolic tangent activation function, was trained on 70% of the data and tested on the remaining 30%. Additionally, the study evaluated the decision tree model's performance through split-sample validation and 10-fold cross-validation. **RESULTS**: The ANN model demonstrated an accuracy of 88.10% in predicting heart disease (AUROC = 0.942) based on six factors and five covariates. The cross-validated tree model achieved an overall predictive accuracy of 84.3%. In contrast, the split-sample validated tree model, which used a balanced 50-50 data split for training and testing, attained an accuracy of 82.0%. OldPeak (ST depression induced by exercise relative to rest) had the highest normalized importance calculated from the MLP ANN model. **CONCLUSION**: Machine learning predictions have gained importance in healthcare, presenting potential benefits in early detection and intervention, leading to improved patient outcomes and reduced healthcare expenses. The study revealed that ANNs outperform decision tree models in accuracy for the dataset in use. Furthermore, the research emphasized the significance of the clinical feature "Oldpeak" in predicting heart failure through ANNs. ANNs can discern intricate relationships between variables and recognize non-linear interactions, a capability sometimes missed by decision tree models. However, the efficacy of machine learning models remains dependent on the quality and volume of the available data.

Table. Model Outcomes.

Model/Parameter	Description/Outcome
Dataset Size	*918 patients*
Training Set	*616 patients*
Test Set	*302 patients*
ANN Model	
Sum of Squares Error (Training)	*57.171*
Sum of Squares Error (Test)	*27.944*
Correct Predictions (Healthy)	*82.70%*
Correct Predictions (Diseased)	*92.60%*
Overall Accuracy	*88.10%*
AUROC	***0.942***
Cross-Validated Tree Model	
Overall Accuracy	*84.30%*
Independent Variables Included	*ST_Slope, ChestPainType, RestingECG, Oldpeak, MaxHR*
Split-Sample Tree Model	
Overall Accuracy	*82.00%*
Independent Variables Included	*ST_Slope, ChestPainType, ExerciseAngina*

Key words: Cardiovascular Diseases; Heart Diseases; Heart Failure; Artificial Intelligence; Machine Learning; Neural Networks; Computer; Decision Trees; Risk Factors (Source: MeSH-NLM).