Title: Machine Learning Classifications of Coronary Artery Disease

Paper Link: https://ieeexplore.ieee.org/document/8692942

1 Summary

1.1 Motivation/purpose/aims/hypothesis

To accurately diagnose coronary artery disease (CAD) using machine learning classifiers.To select the most relevant input features and tune classifier parameters to improve diagnosis

1.2 Contribution

Applied 3 feature selection methods to select 7 most relevant input features. Tuned and tested 3 classification algorithms - SVM, Naive Bayes and KNN. Naive Bayes outperformed the other two with 84% accuracy.

1.3 Methodology

Used Cleveland heart disease dataset with 13 features and 297 entries. Selected features using Information Gain, Correlation and Naive Bayes Classifier Subset. Tuned hyperparameters of classifiers using grid search. Evaluated models using accuracy, recall, specificity and precision

1.4 Conclusion

Naive Bayes gave best performance on the dataset among the tested models. Achieved 84% accuracy in diagnosing CAD.

2 Limitations

2.1 First Limitation/Critique

Single dataset limits generalizability. Only the Cleveland heart disease dataset was used for training and evaluation. Performance may vary significantly on other datasets. Testing on more datasets would better validate robustness

2.2 Second Limitation/Critique

No comparison to other models. Many other classification algorithms exist besides SVM, Naive Bayes and KNN. Advanced algorithms like neural networks and ensemble methods may perform better. Testing other models could identify the optimal approach.

3. Synthesis

The paper demonstrates a systematic approach to applying machine learning for CAD diagnosis. The methodology of feature selection and model tuning provides good performance. Additional datasets and more extensive tuning could further validate the robustness. Overall, it provides a good foundation for using Naive Bayes classifiers for this application.