Paper ID: 1570940718

Paper Title: Principal Component Analysis and Machine Learning in Heart Disease Diagnosis: A Focus on Heart SPECT Parameter Reduction

Authors: Nichapa Srisaranon, Panpatchanan Yimluean and Supisara Bawornwongsatien (Satriwithaya School, Thailand); Piyaon Sapphaphab and Oraya Rinthon (Satriwithaya School, Bangkok, Thailand); Suejit Pechprasarn (Rangsit University, Thailand)

Email: 50120@satriwit.ac.th

Abstract

Single Photon Emission Computed Tomography (Heart SPECT) is a nuclear imaging technique used to assess cardiac function and is integral in diagnosing conditions such as coronary artery disease, heart attacks, and heart failure. Our study utilizes an open-source dataset from the UCI Machine Learning Repository, consisting of 267 instances and 23 attributes (22 predictors and 1 label). After preprocessing, which involved the transformation of biased data to unbiased data, the performance of 32 machine-learning models was assessed. Four models-Kernel Naive Bayes, Coarse Tree, Gaussian SVM, and SVM Kernel-displayed outstanding accuracy, with Coarse Tree, SVM Kernel, and Gaussian SVM achieving the highest accuracy of 80%, and Kernel Naive Bayes obtaining an accuracy of 73.3% (The model achieved the highest accuracy of 80% in training). Subsequently, the principal component analysis identified the least number of variables that could be used: 10 features. Retraining the top four models on these 10 features (F14, F1, F12, F11, F7, F22, F20, F4, F21, F16), resulted in maintained accuracy levels-75% for Kernel Naive Bayes and Medium Gaussian SVM, 76.2% for Coarse Tree, and 72.5% for SVM Kernel. Consequently, the Coarse Tree model, requiring only 10 features, is suggested as the best-reduced model for Heart SPECT diagnosis.