Hierarchical versus Flat Classification of Heart Disease Severity

Francis Biney, *Department of Mathematical Sciences, University of Texas at El Paso*

Maria C. Mariani, *Department of Mathematical Sciences and Computational Science Program, University of Texas at El Paso*

Osei K. Tweneboah, *School of Theoretical and Applied Science, Ramapo College of New Jersey*

Prince Appiah, *Department of Mathematical Sciences, University of Texas at El Paso*

ABSTRACT

In this study, we investigated the predictive performance of two Random Forest approaches for classifying heart disease severity. The models considered were a flat Multi-class Classification (MC) model and a Hierarchical Multi-class Classification (HMC) model. We used the UCI Cleveland Heart Disease dataset and retained the five response classes as an ordinal severity proxy. The MC model predicts all five classes directly. The HMC model first predicts disease presence and then classifies disease-positive observations into four severity groups. Both models were evaluated under eight conditions, including the original data without resampling and seven oversampling, undersampling, and hybrid methods. We used stratified 10-fold cross-validation and compared precision, recall, and F1-score for each class.

The results show that HMC gave higher average scores for Class 2 (Moderate Severity) under all eight conditions. HMC also performed slightly better for Class 0 (No Heart Disease). The combination of HMC and SMOTETomek gave the preferred balance of Class 0 and Class 2 performance, with approximate average scores of 0.91 and 0.69, respectively. However, the flat model performed better for Class 3, and both models gave scores close to zero for Class 4. Errors at the first level of the hierarchy may contribute to the weak Class 3 performance, although the reported summaries do not establish where these errors occurred. Our results support the hierarchical approach when the objective is to improve Class 0 and Class 2 prediction in this dataset. Further work is needed to improve the detection of the higher-severity classes and to evaluate the models on independent clinical data.

Keywords: heart disease severity classification, Random Forest, multi-class classification, hierarchical classification, class imbalance, resampling, SMOTETomek