Abstract

Cardiac arrhythmias are a significant health concern that requires precise and timely detection to prevent severe outcomes, including sudden cardiac arrest. Machine learning methods have gained popularity for automating arrhythmia detection from electrocardiogram (ECG) signals. This study evaluates three widely used algorithms—Random Forest, Support Vector Machine (SVM), and Neural Network—on an imbalanced ECG dataset with five arrhythmia classes. The dataset comprises 175,729 samples, with 87.36% of the data belonging to the normal (N) class, while other classes like F and Q represent less than 0.2%. Preprocessing steps include normalization, feature selection, and dimensionality reduction through *Principal Component Analysis* (PCA). Class-weighted models address the imbalance, and performance is assessed using precision, recall, and F1-score. Results show that the Neural Network model achieved the best overall performance with an F1-score of 99.51%, followed by Random Forest (99.31%) and SVM (98.95%). While these models effectively classified the majority class, they struggled with rare classes like F and Q, highlighting the necessity for advanced techniques to address extreme class imbalance. This research provides insights into the strengths and limitations of these algorithms for arrhythmia classification, offering a foundation for future work in this domain.

Keywords: Arrhythmia, Electrocardiogram, Random Forest, Support Vector Machine, Neural Network, *Principal Component Analysis*