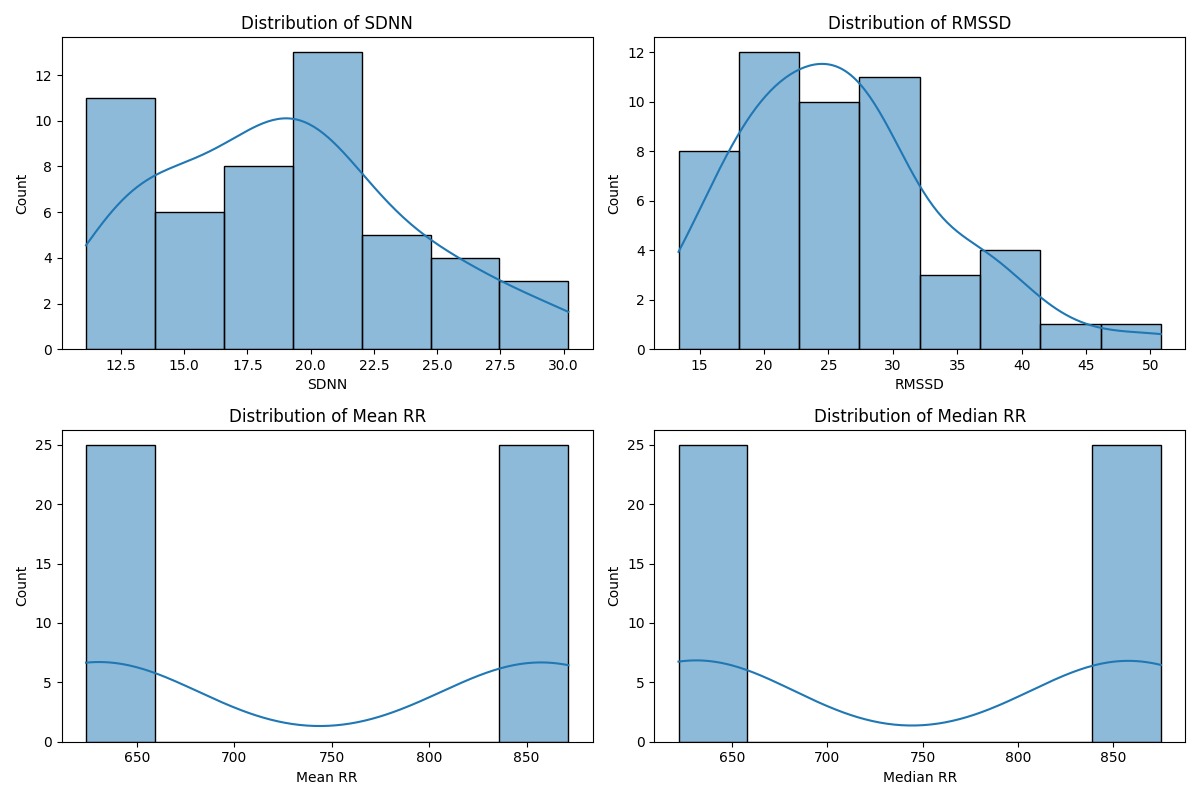
**RESULTS**

**FEATURE DISTRIBUTION**

Distribution of HRV parameter is illustrated in Fig. 1 the distribution of extracted ECG features, including RMSSD, SDNN, Mean RR, and Median RR. These features were derived from the preprocessed ECG signals and serve as key indicators for assessing cardiovascular autonomic function. The distribution analysis helps in understanding the variance and spread of these features across normal and CAN-affected subjects, providing insights into their discriminative ability for classification. Visualizing the spread and central tendency of features like RMSSD, SDNN, Mean RR, and Median RR allows for better understanding of their significance in detecting CAN. By comparing the distributions, we can determine whether certain features effectively separate normal and abnormal cases.



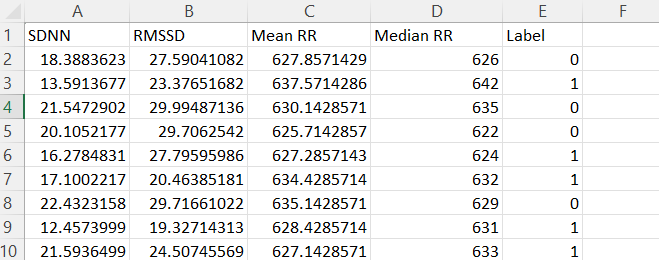
**Fig.1 Distribution of HRV parameter**

**LABELLING THE FEATURES**

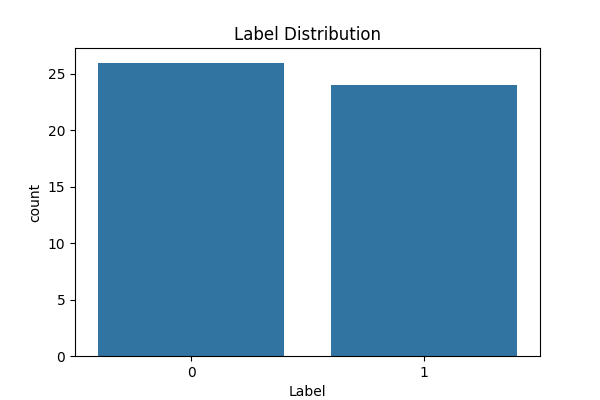
From the extracted ECG features, 10 (20% of the data) sample rows taken were used for classifying normal and CAN cases, and are presented in the Fig.2. These values were obtained through a series of preprocessing steps, including bandpass filtering of the ECG signal, baseline correction, and R-peak detection. The dataset was then labelled based on predefined threshold values for SDNN and RMSSD, where

• 0 indicates normal case

• 1 indicates CAN detection



**Fig.2 Extracted and Labelled CSV file**

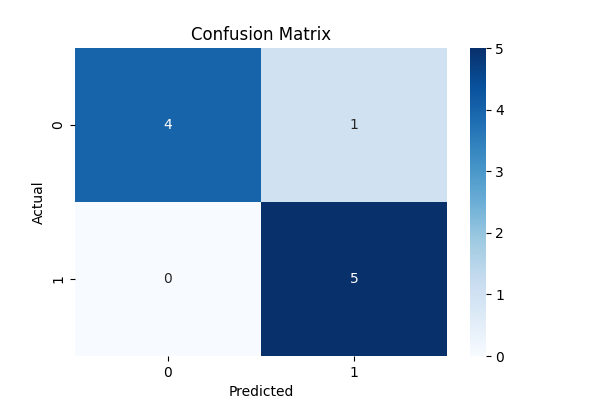
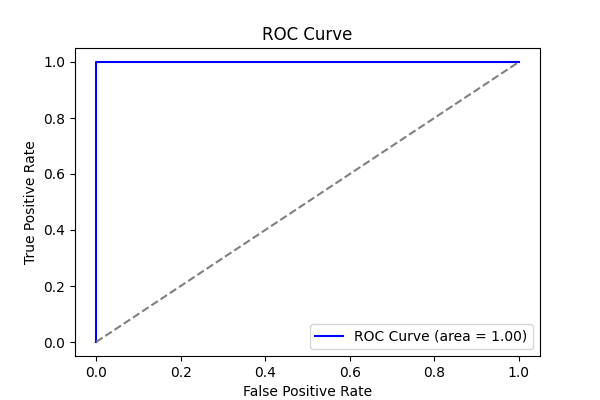
****

**Fig.3 Label Distribution**

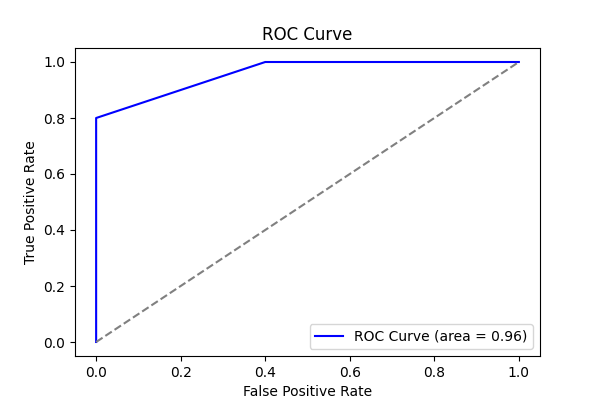
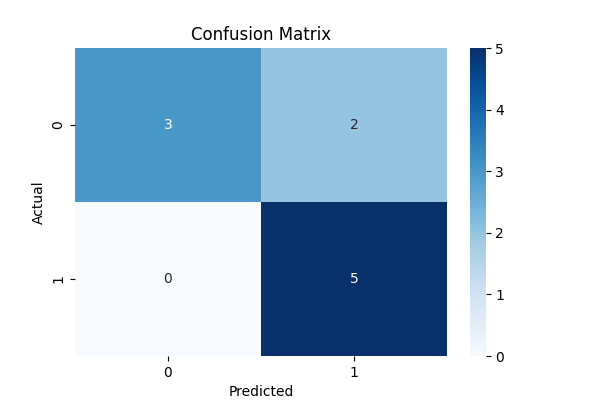
The label distribution graph (Fig.3) provides an overview of the class balance in the dataset. It shows the proportion of normal and CAN-affected cases, which is crucial for assessing model performance.

**CONFUSION MATRIX AND ROC CURVE**

The confusion matrices, as depicted in Fig.4(a, b), offer a comprehensive evaluation of the classification performance of the two models - Random Forest, and XGBoost, by displaying the number of correctly and incorrectly classified instances for both normal and CAN-affected individuals. These matrices are essential in identifying the distribution of errors across different models, allowing for a deeper analysis of their strengths and limitations.

****

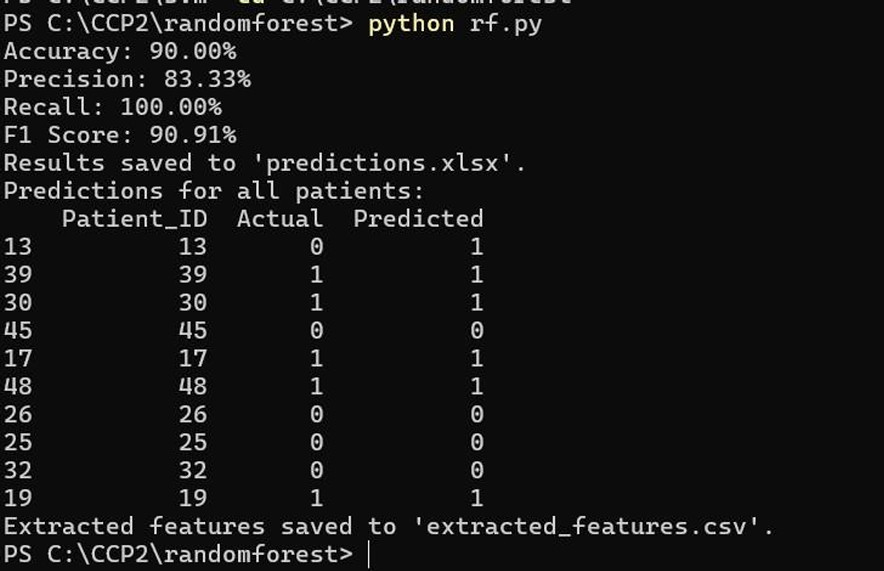
**Fig.4 (a) Confusion matrix, ROC curve of Random Forest**

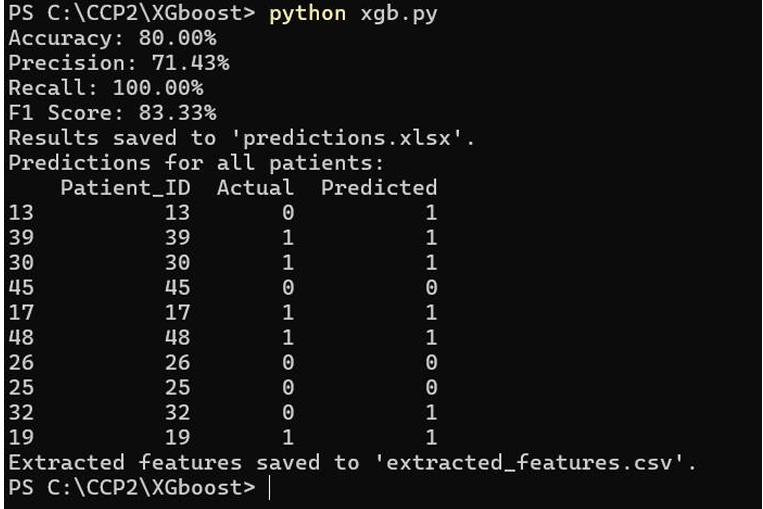
****

**Fig.4 (b) Confusion matrix, ROC curve of XGBoost**

**TABLE.1 PERFORMANCE METRICS**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | Accuracy (%) | Precision (%) | Recall (%) | F1-Score (%) |
| Random Forest | 90.00 | 83.33 | 100.00 | 90.91 |
| XGBoost | 80.00 | 71.43 | 100.00 | 83.33 |

****

****

**Fig.5 Performance Metrics of Random forest, XGBoost**