XAI Project
Multi Label Classification of ECG and SHAP Analysis
DA2 Write UP

Name: R Achuthan

Roll no:21BAI1449

Slot:F1

1. Data Preparation and Preprocessing:
- Initially, we received a dataset consisting of ECG signals stored in various files.
- Each ECG signal file represents a specific heart condition, categorized into different subgroups such as NOR, LBB, RBB, PVC, and APC.
- We segmented the ECG signals into 10-second intervals and converted them into spectrogram images using libraries like `wfdb` and `scipy`.
- The spectrogram images were saved in respective folders based on their heart condition subgroup.

2. Dataset Splitting and Organization:
- We split the dataset into training and testing subsets according to the specified ratios for each heart condition subgroup.
- The training and testing subsets were organized into separate folders named "train" and "test" within a parent "data" directory.
- Each folder within the "train" and "test" directories represents a heart condition subgroup, containing the respective spectrogram images.

3. Neural Network Model Creation:
- We designed a convolutional neural network (CNN) model using Keras/TensorFlow for multilabel classification of heart conditions based on spectrogram images.
- The CNN architecture comprised multiple convolutional layers followed by max-pooling layers, dropout layers, and dense layers with appropriate activation functions.
- The model was compiled using categorical cross-entropy loss and accuracy metrics, with the Adam optimizer.

4. Model Training and Evaluation:
- We trained the CNN model using the spectrogram images from the training dataset.
- Training progress and model performance were monitored using validation data to avoid overfitting.

- Evaluation metrics such as accuracy, loss, and confusion matrix were computed to assess the model's performance on the testing dataset.

5. SHAP (SHapley Additive exPlanations) Analysis:
- We used SHAP to interpret the predictions made by the CNN model on individual spectrogram images.
- SHAP values were computed to quantify the impact of each pixel in the spectrogram on the model's output.
- Visualization techniques were applied to visualize the SHAP values and gain insights into the model's decision-making process.

6. Future Directions:
- Further optimization and fine-tuning of the CNN model could be explored to improve classification accuracy.
- Additional interpretability techniques and visualization methods could be employed to gain deeper insights into the model's behavior and decision boundaries.