Cardiology Department Detailed Overview

The Cardiology Department focuses on predicting the risk of heart diseases, specifically heart attacks, and classifying ECG images into multiple classes. Here's a detailed breakdown of the work done in this department:

ECG Image Classification (6 Classes)

Initial Approach:

Initially, we worked with tabular data for ECG, which included 187 columns representing features, and the 188th column representing the class labels (4 classes).

The model achieved 97.5% accuracy, but we wanted to improve the classification by extending the analysis to more classes.

Improvement Using ECG Images:

We shifted to using ECG images, specifically focusing on Lead 2 ECG images.

We converted these images into a DataFrame, where each row contained the image path and the classification label.

The dataset was split into train and test, with each dataset organized into 6 folders, where each folder contained images of one class.

Data exploration and analysis (EDA) were performed on both the training and test DataFrames.

Data Imbalance:

A significant issue we encountered was the imbalance in the dataset. The smallest class had fewer than 1,000 images, while the majority class had up to 10,000 images.

To address this, data augmentation techniques were applied to balance the dataset, and class weights were used during model training. This led to high train and test accuracy, reaching 99%.

Confusion Matrix:

Despite the high accuracy, the confusion matrix revealed that the model performed poorly on the minority classes, with accuracies for these classes dropping below 20%, except for the majority class.

To handle the imbalance, I applied down-sampling to the majority class (limiting it to 8,000 samples) and over-sampling to the minority classes.

Modeling with VGG16:

After re-sizing the ECG images to 224x224 pixels, I applied the VGG16 model.

Initially, the model achieved 95% train accuracy but suffered from overfitting with a high gap between train and test accuracy.

To resolve over-fitting, I reduced the model complexity, removed some layers, and adjusted the dropout rates.

Final Model:

After these adjustments, the model surprisingly achieved 100% test accuracy, and a further check using the confusion matrix showed perfect performance across all classes, with precision, recall, and F1 score all being 100% for every class.

This result was significant, indicating that the model had learned the underlying patterns well without over-fitting.

Heart Attack Prediction

Dataset:

The department focused on predicting the likelihood of a heart attack using patient data such as , blood pressure, cholesterol levels, etc.

Data Preprocessing:

Data Scaling was applied to ensure that the features were on a similar scale, which helped improve model performance and stability.

Model Comparison:

Various models, including XGBoost, Random Forest, and Decision Tree, were trained and evaluated.

XGBoost was ultimately chosen because it provided the highest accuracy, precision, and recall, achieving 98.4% in all metrics.

Ensemble Modeling:

Although XGBoost was selected, an ensemble modeling approach was initially considered for boosting prediction performance. However, XGBoost proved to be the most effective model.

Deployment:

For the deployment phase, data visualization was integrated to enhance the clarity of predictions.

A waterfall chart was used to display the importance of each feature in the heart attack prediction.

A gauge chart was implemented to visually represent the probability of heart attack risk, providing an intuitive view for doctors.

Summary

In the Cardiology Department, a thorough and systematic approach was used to improve ECG image classification by addressing data imbalance, applying advanced augmentation techniques, and leveraging powerful deep learning models like VGG16. Additionally, for predicting heart attack risk, ensemble learning methods were employed to combine multiple models, ensuring high accuracy and reliable predictions. This work is critical, as accurate heart disease risk assessment can significantly improve patient outcomes and prevent life-threatening conditions