CSCI 5980/8980: Machine Learning for Healthcare
Final Project Presentation - Spring 2024

Arrhythmia classification using ML/DL

Group 7:
Anwesha Samaddar, Hemanth Kumar Tirupati,
Manpreet Singh, Sharan Rajamanoharan

Intro/abstract

In this project, we assess various ML/DL Approaches for Cardiac Arrhythmia classification on the MIT-BIH dataset. Our objective is to offer an effective and time-saving ECG analysis solution by combining signal processing techniques with efficient deep learning models.

Main Contributions of the Project

- Experiment with various Neural Networks for ECG interpretation.
- Classified a broad range of arrhythmias into 5 distinct classes.
- Compared classification performance across these approaches.

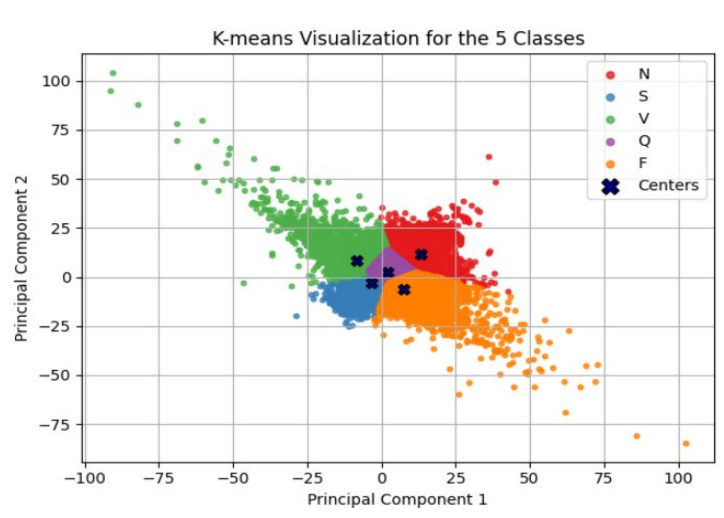
Dataset & Preprocessing

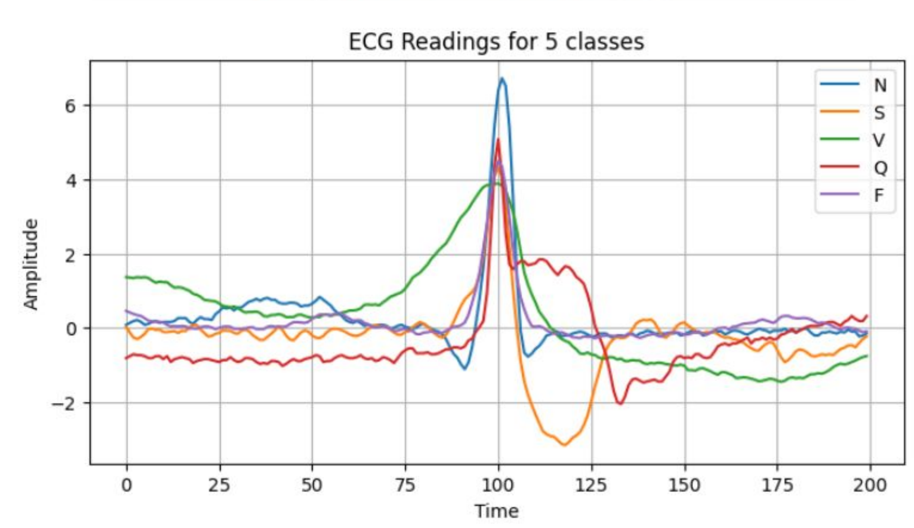
Our dataset is MIT-BIH dataset, which contains **48 ECG** recordings. These recordings are **30 min** long, with sampling frequency of **360 Hz**. We divide our dataset into 200 samples length, with each sample centered at QRS detected waves. This makes our total data size **107400 by 200**. Additionally, we store a label column corresponding to each row making data shape **107400 by 201.**

Preprocessing:

- Next, we scale and segment signal to 5 classes.
- 5 arrhythmia Classes: **N**(normal), **F**(fusion), **Q**(unknown) **V**(ventricular ectopic), **S**(supraventricular ectopic)

```
dic = {'N':['N','L','R','B'],
       'S':['A','a','j','S','e','j','n']
       'V':['V','r','E'],
       'F':['F'],
       'Q':['Q','?','f','/'],}
```





Results & Discussion

- We are comparing the test set accuracy, validation loss and average F1-score (across 5 classes) - for Random forest, MLP, CNN, LSTM and CNN-LSTM.

Model Name	Trainable parameters	Model Test Accuracy	Validation Loss	F1 (weighted average)
Random Forest	100 trees	98.05%	----	0.9791
MLP	26,555	97.97%	0.0650	0.9782
CNN	2,470,269	98.66%	0.0647	0.9861
LSTM	72,205	98.51%	0.0647	0.9848
CNN-LSTM	840,653	98.38%	0.0472	0.9832

Discussion

We can see that CNN performs the best based on test accuracy and confusion-matrix however it also has 2.4 Million parameters. While CNN-LSTM, CNN and LSTM have similar performance, LSTM had significantly fewer trainable parameters, least training time and smallest model size. LSTM being lightweight can be preferred for real-time arrhythmia diagnosis.

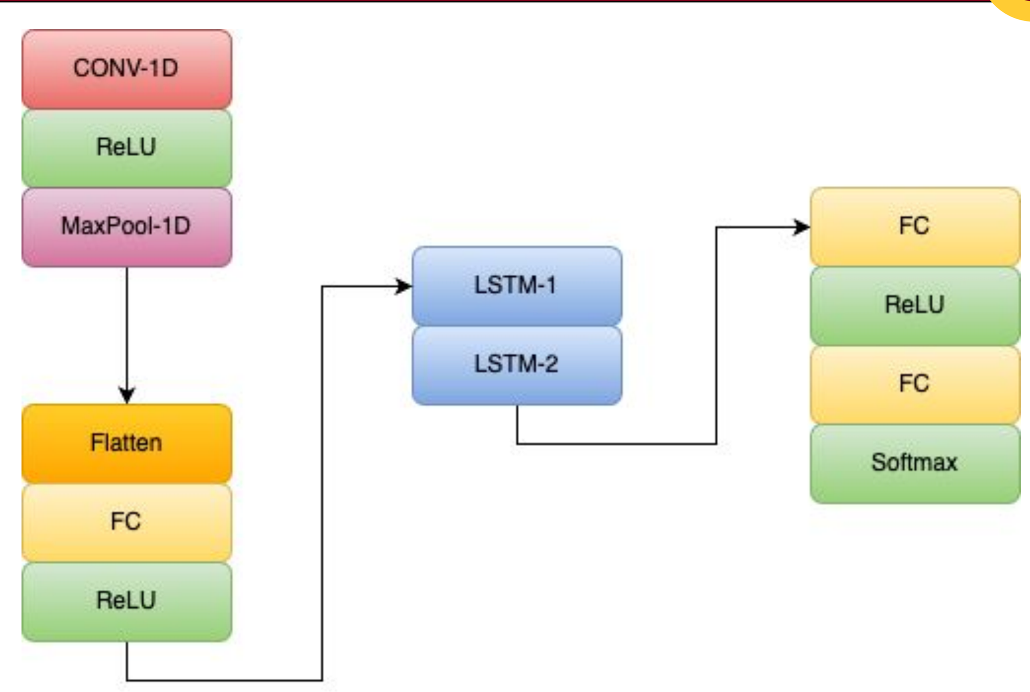
Background

An Electrocardiogram (ECG) is a standard diagnostic tool used in healthcare that tracks the ***electrical activity of the heart*** over time to assess cardiovascular health. **Arrhythmia categorization** using ECG monitoring is critical for predicting a variety of cardiac problems, including cardiovascular illness, stroke, and atrial fibrillation identification, and hence contributes greatly to disease prediction and treatment.

Current ECG analysis relies on manual methods, which are time-consuming and require expertise. This has motivated the development of ML/DL approaches to provide efficient, cost-effective solutions for arrhythmia diagnostics.

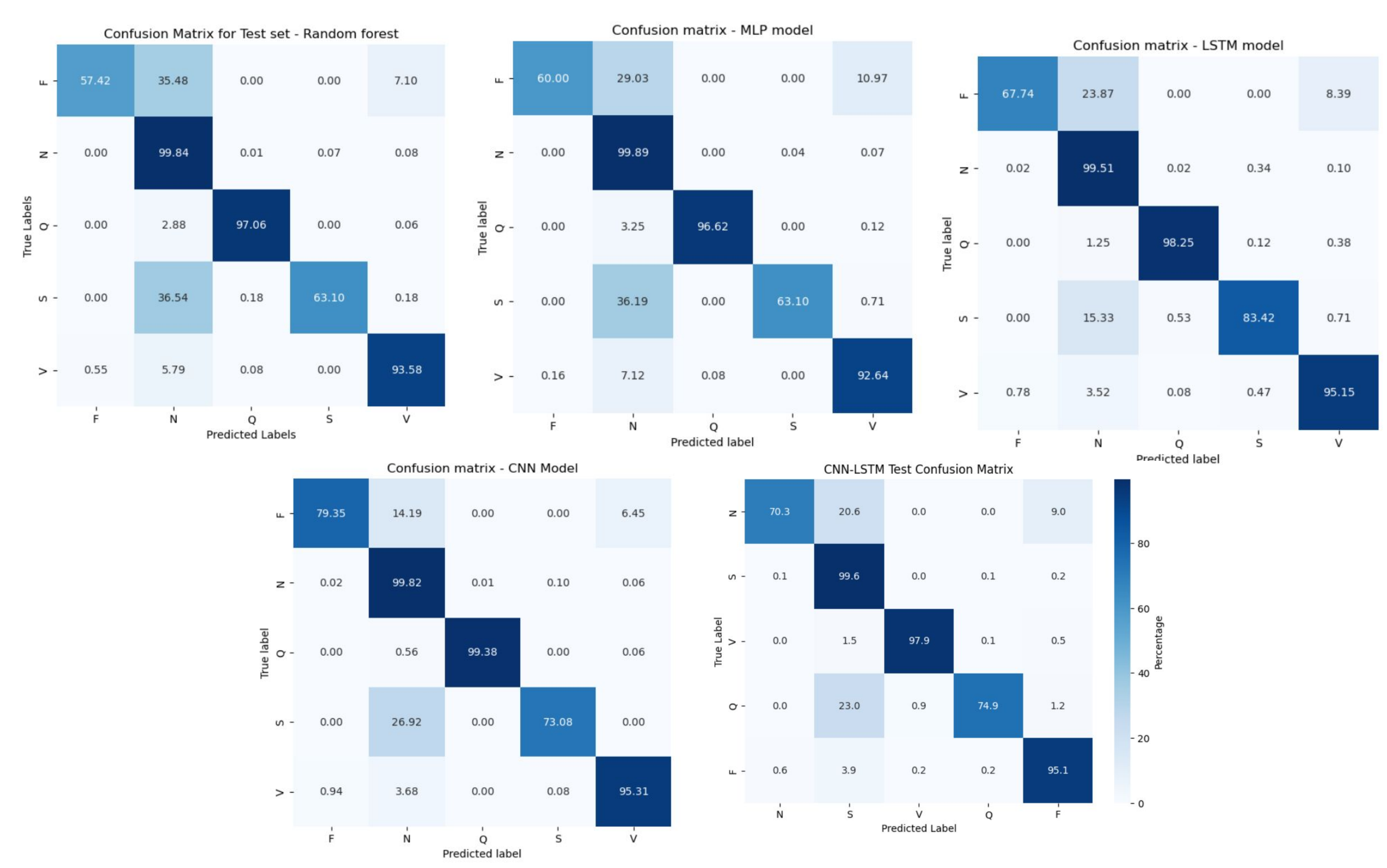
Proposed Approach & Evaluation

We tried out a CNN, LSTM and a combination CNN-LSTM network (shown in adjacent figure) and compared the classification performance.



Evaluation

- We compare the confusion matrices for each of our five models to get a detailed breakdown of the performances. It shows us the percentage of accurately classified samples and misclassifications with counts of true and predicted labels.



Conclusion & Future Directions

Conclusions

Arrhythmia Detection is a well-studied problem and the accuracy of the current methods is encouraging for clinical use. Deployment of such AI systems in rural areas world greatly increase access, reduce costs and addresses medical personnel shortage.

Future Directions

- Potential Future Experiments in Arrhythmia Detection include trying out other Neural Network Architectures like Bidirectional-LSTM and Transformers.
- We could also experiment on datasets like PTB-XL , CinC and CPSC .

References

1. Cardiologist-Level Arrhythmia Detection with Convolutional Neural Networks
2. Classification of Arrhythmia in Heartbeat Detection Using Deep Learning
3. A Literature Review: ECG-Based Models for Arrhythmia Diagnosis
4. MIT-BIH Arrhythmia Database