ECG Heartbeat Classification

Practical 1 - Report

Chu Hoang Viet - Data Science

1 Introduction

This is a practical work 1 report of Machine Learning in Medicine course. In this report, I will analyze the MIT-BIH Arrhythmia dataset and propose a model for the classification task.

2 Dataset

The MIT-BIH Arrhythmia contains 187 features that represent the heartbeat in 125 Hz sampling rate. There are 5 classes in this dataset: Normal, Atrial premature, Premature ventricular contraction, Fusion of ventricular and normal, Paced. The training set includes 87,554 samples, while the testing set has 21,892 samples. However, both training and testing data are imbalanced, with many samples in the Normal class.
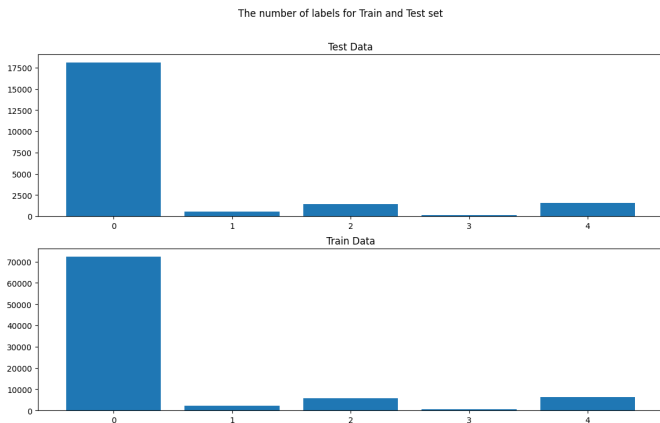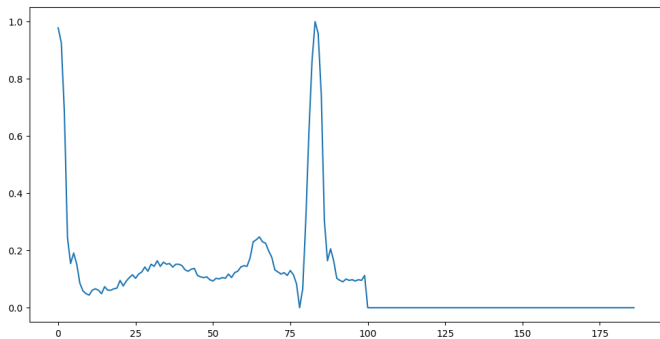

Figure 1: The number of classes for Train and Test set

Figure 2: Heartbeat of 1 sample

3 Model

I try to use the similar method as the original paper[1] but less complexity by reducing the number of layers.
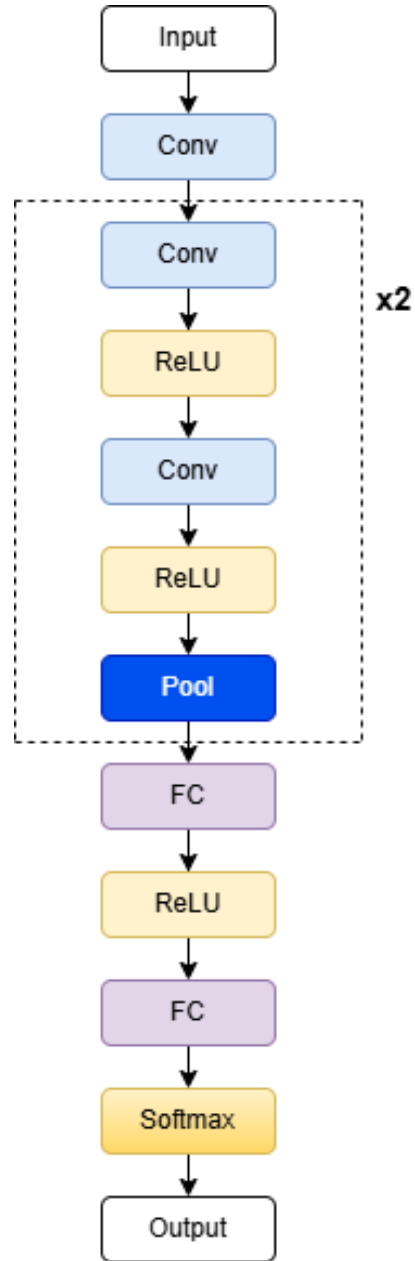

Figure 3: Model pipeline

The **convolutional block** is used to extract features and then it will be fed to the **fully connected layer** to learn the pattern. **Softmax** function takes responsibility for predicting the probability of each class.

4 Result

Due to the data has not been pre-processed, the training loss is fluctuated between 2 values: 1.40483 and 0.90483. After applying model on the test set, it gives the accuracy of 82.76%, which is smaller compared to 95.9% of original paper. The imbalanced dataset could be a cause for this

problem, the model has learned more on this overweighted class.

References

[1] Mohammad Kachuee, Shayan Fazeli, and Majid Sarrafzadeh. "ECG Heartbeat Classification: A Deep Transferable Representation". In: *2018 IEEE International Conference on Healthcare Informatics (ICHI)*. IEEE, June 2018, pp. 443–444. DOI: 10.1109/ichi.2018.00092. URL: http://dx.doi.org/10.1109/ICHI.2018.00092.