Shiraz University

Neural Networks and Deep Learning (Mordad 1401)

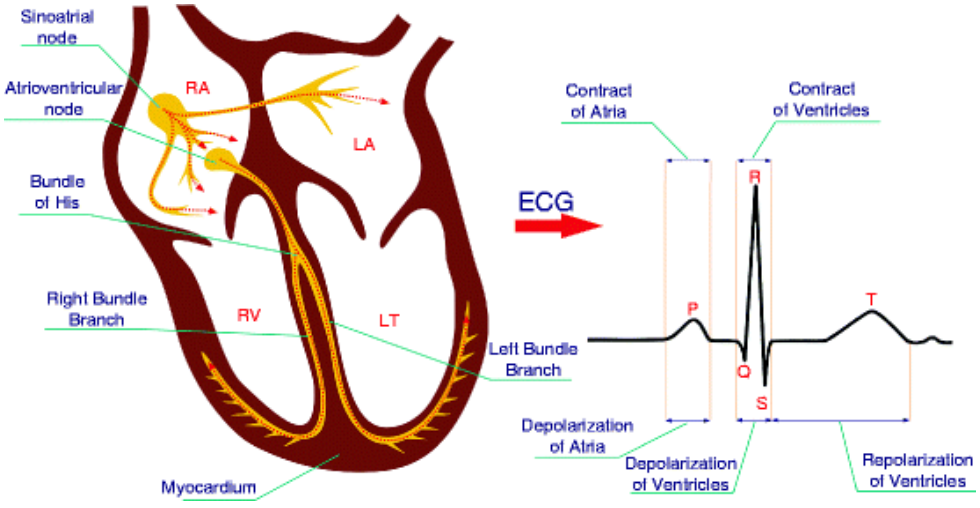
Assignment #4: Detection and Identification of Electrocardiogram Signals using Long Short-Term Memory (LSTM)

Due date: 21st Mordad 1401

To do this homework, go through theories and concepts from Long Short-Term Memory (LSTM). If you're coding in Python or Matlab, you can use any library for LSTM Networks.

Electrocardiogram :

An electrocardiogram records the electrical activity of the heart at each contraction. When an electric wave is generated in the heart, the inside of the heart cell quickly becomes positive relative to the outside. Stimulation by an electric wave nourishes the polarity of the cell.

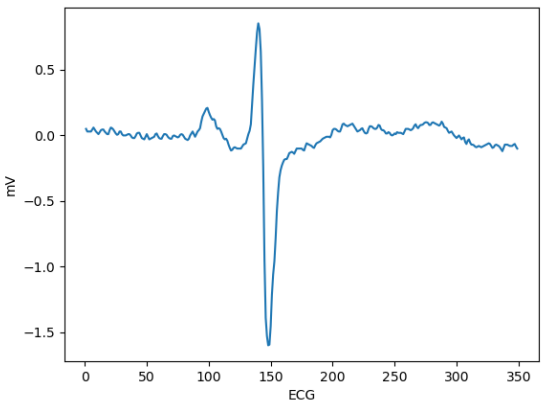


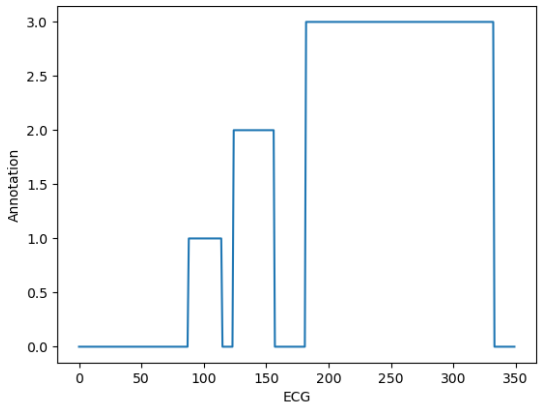
The most important characteristic ECG signal that cardiologists in the diagnosis of heart disease employ, the QRS complex. This characteristic is more important than T, P waves, and other signal properties because it is

easier to distinguish and separate from the ECG signal than other features. Also, the QRS complex shows ventricular depolarization, which plays the most important role in the electrical activity of the heart. Therefore, the diagnosis and isolation of the compound are crucial in the classification and diagnosis of cardiac abnormalities. Also, by diagnosing and counting QRS, the intensity of the heartbeat and its possible inconsistencies can be observed and examined.

Dataset: QT database

In this assignment, you will predict the secondary sequences (annotations) from their primary sequences (samples). You are given a dataset of 100 files. Each sample file contains a primary sequence, ‘mV’ of signal, and its secondary is in annotation file ‘annotation’. Each primary and secondary sequence contains data streams of continuous values and 0 to 3, respectively (1 for P wave, 2 for QRS complex, 3 for T wave, and 0 for other points of signal). Thus, you have 60 input and output patterns of data streams for a training dataset. An example of input and output data streams is shown below.





Your Task:

1. Divide each file (with 15000 or 2500 samples) into records with 500 samples. **(Be careful! Don't shuffle the samples.)**
2. After that, you should have 500 sample records with annotation. Use a sliding window of size 5 to get training patterns for each record.

3. After the training patterns is prepared, build a bidirectional LSTM network for classification and train it. (Present training sequences of each record forward and backward to two separate LSTM layers, both of which are connected to the same output layer)
4. Repeat steps 1 and 2 for the test data and then classify them.
5. Calculate the classification accuracy of the predictions.
6. Use a sliding window of sizes 10 and 15 for each sequence and repeat steps 3 to 5.

Notes:

- **Pay extra attention to the due date. It will not extend.**
- **Be advised that submissions after the deadline would not grade.**
- **Prepare your full report in PDF format and include the figures and results.**
- **You can use any library in Matlab or Python.**
- **Submit your assignment using a zipped file.**

"StdNum_FirstName_LastName.zip" to soroushmehrpou@gmail.com.