ECG Signal Classification Using Neural Network

Yashu Vashishath, Tallon Coxe, Boyi Zheng, Jongwook Yoon, Ambarish Kumar

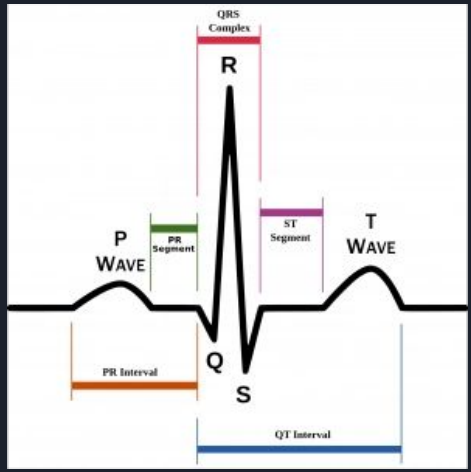
Problem

What is the ECG signal?

- Electrocardiography is the process of producing an electrocardiogram, a recording of the heart's electrical activity.

How to diagnose heart problems using the ECG signal?

- ECG signal consists of different waves named as PQRST waves.
- Cardiologists diagnose cardiac abnormalities using interval of peaks and amplitude of peaks.



https://en.wikipedia.org/wiki/Electrocardiography

Dataset

How to pre-process the PTB-XL (ECG dataset)?

- Filtering noises:
  Baseline wander, Powerline Interference, and electromyographic (EMG) noise.
- Extracting features (interval of peaks and amplitude of peaks) of the PQRST wave.

What does the pre-processed dataset look like?

- 18935 records and 148 features including ID, sex, age, and PQRST-related features.

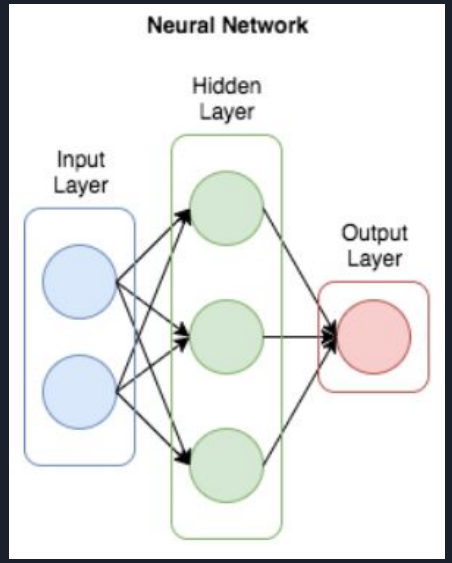
ID	age	sex	P_width_1	P_width_2	P_width_3	P_width_4	P_width_5	P_width_6	P_width_7	...	T_height_ecg_5
0	56.0	1	3.0	6.0	3.0	6.0	7.0	5.0	6.0	...	0.039055
1	19.0	0	7.0	15.0	10.0	4.0	4.0	8.0	5.0	...	0.151352

- 5 diagnostic classes: NORM (1), MI (2), STTC (3), CD (4), HYP (5)

Proposed model

What is the our model?

- Neural Network Model in R.
- Accuracy is used as metric.
- Sigmoid is used as activation function.



https://ebbnflow.tistory.com/119

Results

Comparison of baseline models and the proposed model.

- baseline models: Logistic regression, Support vector machine, Random Forest

	accuracy_train	accuracy_test	accuracy_shuf
Logistic Regression	0.6720223	0.02236703	0.5658683
SVM	0.7110206	0.5658683	0.5658683
Random Forest	0.9948706	0.1521663	0.5658683
NN	0.6444143	0.593695	0.5658683

Discussion

- Due to unbalance data the machine learning models are not performing good.
- Making the models perform better we will need to work more on data preprocessing steps by either making the data balanced (removing normal samples or resampling of samples with different diseases).
- There can be more steps taken to improve the models by tuning the hyperparameters which was not done due to time constrained in the project.

Thank you!