Assignment 3 (100 points)
Heartbeat Sounds - Kaggle Competition

1. **Dataset**
   This dataset was originally for a Machine/Deep learning challenge to classify heartbeat sounds. The data was gathered from two sources: (A), and (B).

   (a) From the general public via the iStethoscope Pro iPhone app.

   (b) From a clinical trial in hospitals using the digital stethoscope DigiScope.

2. **Tasks**
   There were two challenges associated with this competition

   (a) **Heart Sound Segmentation**
   The first challenge is to produce a method that can locate S1(lub) and S2(dub) sounds within audio data, segmenting the Normal audio files in both datasets.

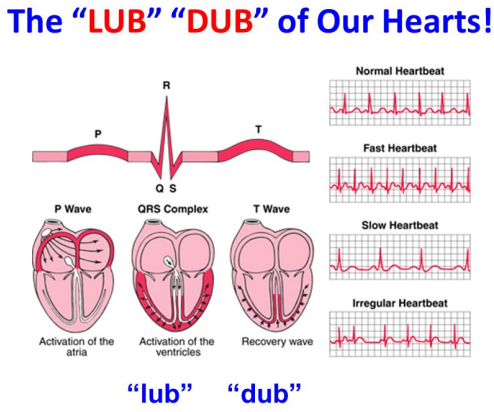

Figure 1

   (b) **Heart Sound Classification**
   The task is to produce a method that can classify real heart audio (also known as "beat classification") into one of four categories.

3. **Data Exploratory analysis** (15 points)
   Perform any data analysis techniques you know (or even learned from your previous exploration) to help in better understanding your data distributions, anomalies, and other possible information related to each task.
   Adding more techniques in this part may help you to achieve a better model performance which will help you to get (10 points Bonus!).

Experiment number	Architecture	LR	Loss	Val f1	test f1
1	CNN	0.001	Cross-Entropy	78	75
2	ANN	0.001	Cross-Entropy	75	70
3	RNN	0.001	Cross-Entropy	80	77

Table 1: Experimental Results

4. **Data splitting** (15 points)

   (a) Perform a splitting with a ratio 70, 15, 15 for each partition and **fix splits for all experiments.**

   (b) Use validation partition for the selection of your best models.

   (c) Reports your models' performance tabulated for each experiment with its hyper-parameter used for training (report test partitions result based on the highest performing model on your validation partition).

   (d) Never uses a test partition for your evaluation of the model while training (this behavior will be penalized), only validation set allowed for -after epoch- evaluation.

   (e) Report one model for each change in your training parameters(loss, Architecture, Learning rate .... etc.).

5. **Experiments setup** (70 points)

   Tasks mainly split into two tasks first one is a regression task second task is a classification task.

   Model Architecture selection (20 points for each task)
   You're required to build 2 types of networks Mainly Feed Forward Neural Networks and Convolution Based Networks (CNNs) adding more experiments with different architectures can help you to gain better performance (not necessary).

   Loss Selection for your task (7.5 points for each task)
   Based on your task you can perform an experimental analysis of multiple losses suitable for your task, reading PyTorch documentation for losses may help you in your selection.

   Evaluation Metric selection (7.5 points for each task)
   You can select evaluation metrics based on your task but mainly f1-score for classification and Mean squared error for regression will be compared for best and second best results (will be given extra 10 points as a bonus).

Notes:

You should provide a detailed report illustrating each step and its findings and opinions, missing parts from the report will be penalized by 1 to 2 points from its total points.

You should provide a table of your best model performance on the test partition in a similar table 1.
Only teams of 1 to 3 allowed.