Introduction:

For this project, we are doing 2 tasks, classification & regression. Classification requires us to train a neural network to classify fetal heart rate (FHR) and uterine contraction (UC) according to 3 classes (Normal, Suspect & Pathologic). The classification in the dataset have been manually done by experts & we are training the model based on this dataset.

For regression, we are training a neural network to predict the chances of a graduate getting into a Master Program. We have parameters such as GRE score, Letter of recommendation, etc that we can use to make our prediction. The dataset also contains the predicted chance of entering a particular university.

**Conclusions**

From the classification task, we find that the optimal decay parameter is 0. This could mean that the decay parameter is not suitable for this task. It could also mean that we were unlucky and our test dataset is unevenly distributed. We believe that we were unlucky as class 3 is quite rare and should have a big impact on accuracies in different datasets. We can observe this could be the case if we compare figure 3c & 4c. The test data has similar shape for both figures. Even if we were to shuffle every time we do k-fold validation, the issue persists as one dataset may end up biased due to luck.

From the regression task, we find that not all parameters are equally important. Some are even detrimental, as their effect on the output is not consistent. Parameters should occupy a bigger input space & binary parameters are less desirable.

Finally, although we have found that 4 layers is the optimal number of layers, we got to this number by trial & error. We do not know the reason why. If we were to make a guess, it is likely that shallow networks have too little space to process parameters while deep networks are difficult to train. We believe that each neural network requires a minimum depth to process all parameters. However, modern neural network techniques are unable to properly backpropagate weights to create deep neural networks effectively.