

# MACHINE LEARNING Assignment 3 Report

110062219

December 5, 2023

## 1 Difficulty Encountered

In the beginning, I implemented focal loss function using just one single line of Python code and got passed the sample test case. Nonetheless, when I started training the model, some runtime error occurred. After some debugging, I correct my implementation.

## 2 Structure of Classifier

### 2.1 Basic

6 hidden layers. The first 5 ones consist of 10 neurons with ReLU activation function, and the last one are 2 neurons with softmax.

### 2.2 Advanced

3 hidden layers comprising 256, 100 & 10 neurons, respectively. The first 2 ones use ReLU as activation function, whereas the last one adopt softmax.

## 3 Effort to Improve Models

Initially, I picked a learning rate that was too slow to converge. Moreover, I had too many layers and too many neurons in each layer for both basic and advanced part, which might led to overfitting.

Last but not least, after applying random mini batch to advanced part, notable improvement was achieved.