

# ML HW3 Report

---

***Describe what problems you encountered and how did you solve them when implementing the basic and advanced functions.***

在 basic part 我用 cross entropy loss 來 train，model 的 error 會越 train 越大，結果是我沒看清楚 CCE 的 Y 是要丟 one-hot encoding 的 array 進去。

在 evaluate basic part 的 model 時我遇到一個問題，我知道在 train model 時，當 validation loss 不降反升時，就應該要停止，以防 overfit，但在這個作業我遇到的情況是，稍微 overfit 的 model 能得到較好的 f1 score，那我應該要選 validation loss 最低的 model 還是 f1 score 最高的 model？

我後來選擇了前者：validation loss 上升時就停止，但我還是不太確定哪個比較好。

***Briefly describe the structure of your binary and multi-class classifiers.***

由於 data class imbalance 的問題，兩個 classifier 我都使用 focal loss，

Binary classifier，我使用以下架構：

```
layers_dims = [10,7,5,2]
activation_fn = ['relu','sigmoid','softmax']
```

在 multi-class 使用以下的 architecture：

```
layers_dims = [28*28, 15*15, 13*13, 7*7, 5*5, 20, 10]
activation_fn = ['relu','relu','relu','relu','relu','softmax']
```

***Describe effort you put to improve your model (e.g., hyperparameter fine tuning, losses' impact on the result).***

1. 用 focal loss，且用 class 的比例來計算適合的 alpha 值
2. stratified\_train\_val\_split，我寫了一個 function 會把 training dataset 切 0.2 出去當 validation dataset，同時讓兩個 dataset 維持一樣的 class distribution，結合第一點，f1 score 有顯著上升
3. 在 basic 與 advanced 我都使用了 early stop 的機制，當 validation loss 上升時就停止，防止 overfit，以此找到 basic part 最佳的 epoch 數。