

110590007 資工系大三 白宸安

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

在實作上遇到的最大問題是在寫 softmax + focal_loss 的 backward，因為公式很複雜，加上要用 code 寫出來，所以實作上遇到很多麻煩，比如 index 對不上、要搞懂 shape 的對應等，debug 了很久，慢慢一行一行看，最後才弄出來。

另外則是在 basic 和 advanced part 的 classifier，在處理資料以及實作 classifier 時也花了不少時間，最大難點就是要對應好 shape 的樣子，我覺得這個是我自己最感到困難的部分，後來也是透過 print 出來看慢慢解析才寫出來。

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

Binary classifier：

1. 首先會先將資料做 split，分成 training 和 validation，接著將資料分布畫出來，以觀察情形。
2. 可使用 sigmoid + cross_entropy 或 softmax + focal_loss，這邊使用後者。
3. 設定好 layers_dims、activation_fn、gamma、alpha 等要用到的參數。
4. 將 y 轉換成 one hot encoding。
5. 接著就使用 learning_rate = 0.01、num_iterations = 12000 開始跑 batch gradient descent training。

Multi-class classifier：

1. 一樣將資料做 split，分成 training 和 validation，接著將資料分布畫出來，以觀察情形。
2. 使用 softmax + focal_loss
3. 設定好 layers_dims、activation_fn、gamma、alpha 等要用到的參數。
4. 將 y 轉換成 one hot encoding（有 10 個 class）。
5. 使用 learning_rate = 0.008、num_iterations = 300 開始跑 mini batch gradient descent training。

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

主要進行了 hyperparameter fine tuning，進行了多次實驗，包括跑了 sigmoid + cross_entropy 或 softmax + focal_loss，調整 learning_rate、num_iterations，還有調整 gamma 和 alpha 等。最後有測試出一些不錯的參數。

此外，有發現大致上 loss 越小會讓 result 越精準，但只要超過一定程度，就會導致 overfitting，會變成預測 training set 很精準，但 validation set 反而變不准的情況。