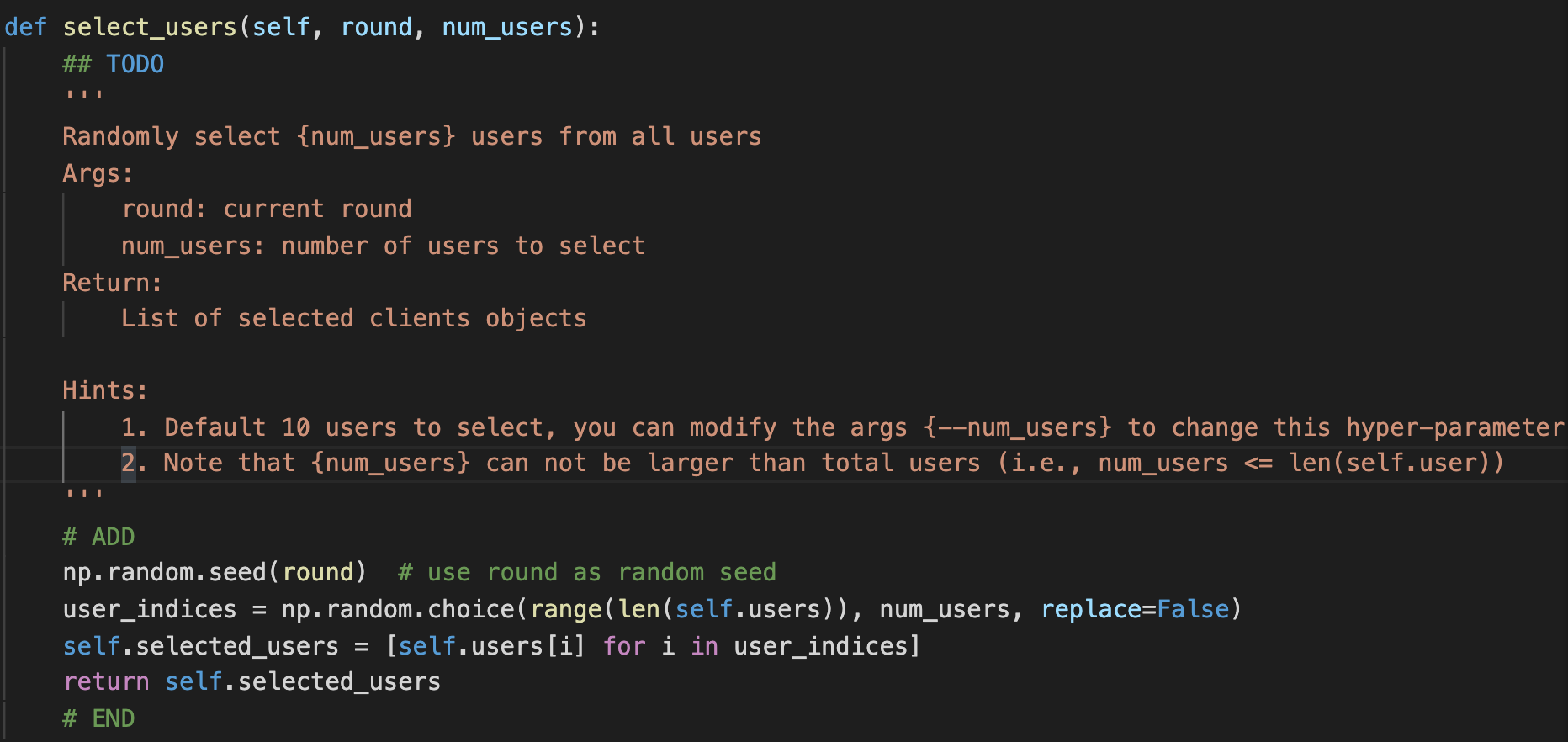
HW5 report

1. Explanation

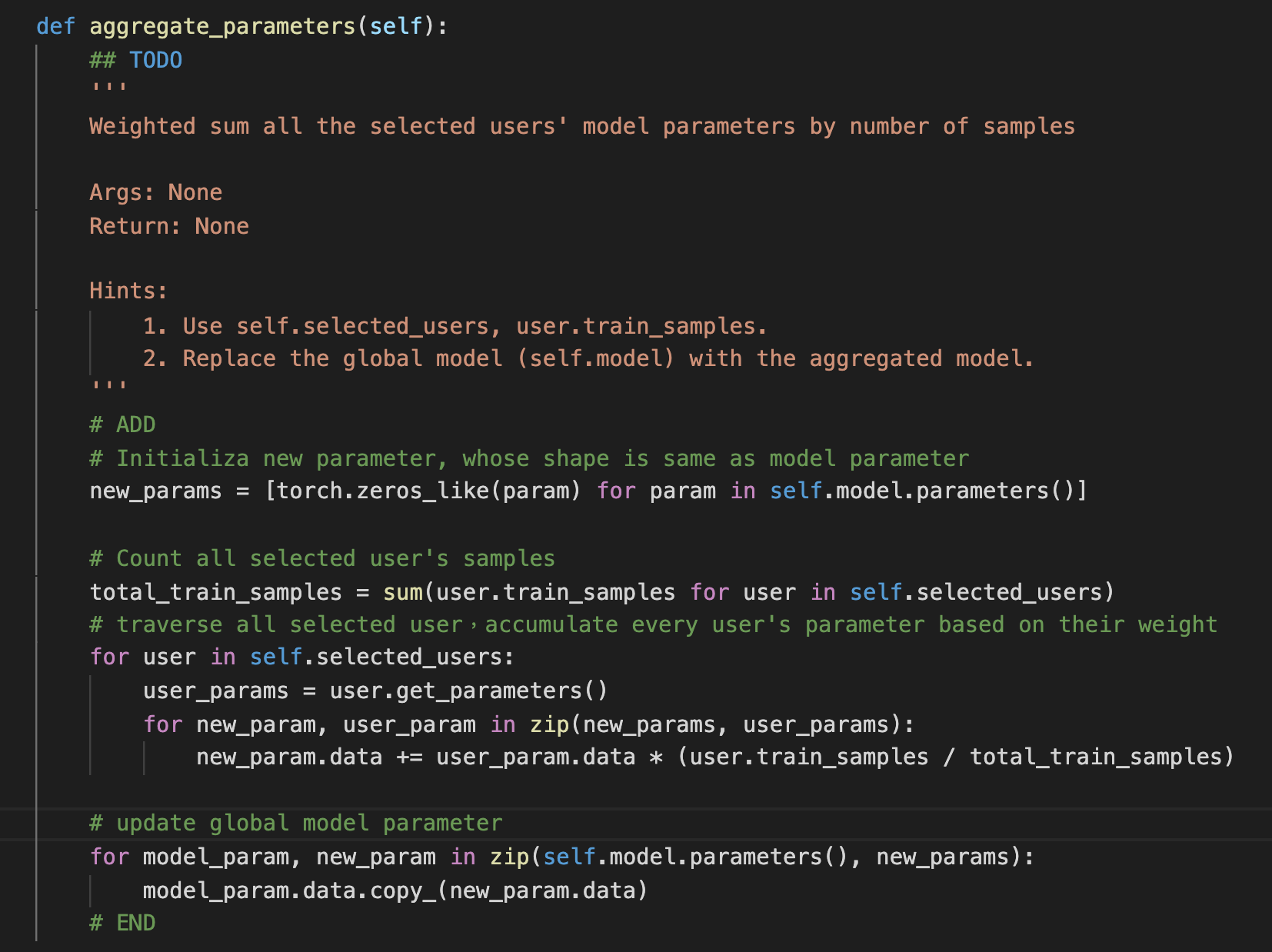
severbase.py

select\_users(self, model, beta)



使用當前round作為random seed，隨機選擇users並且將結果存入self.selected\_users中。

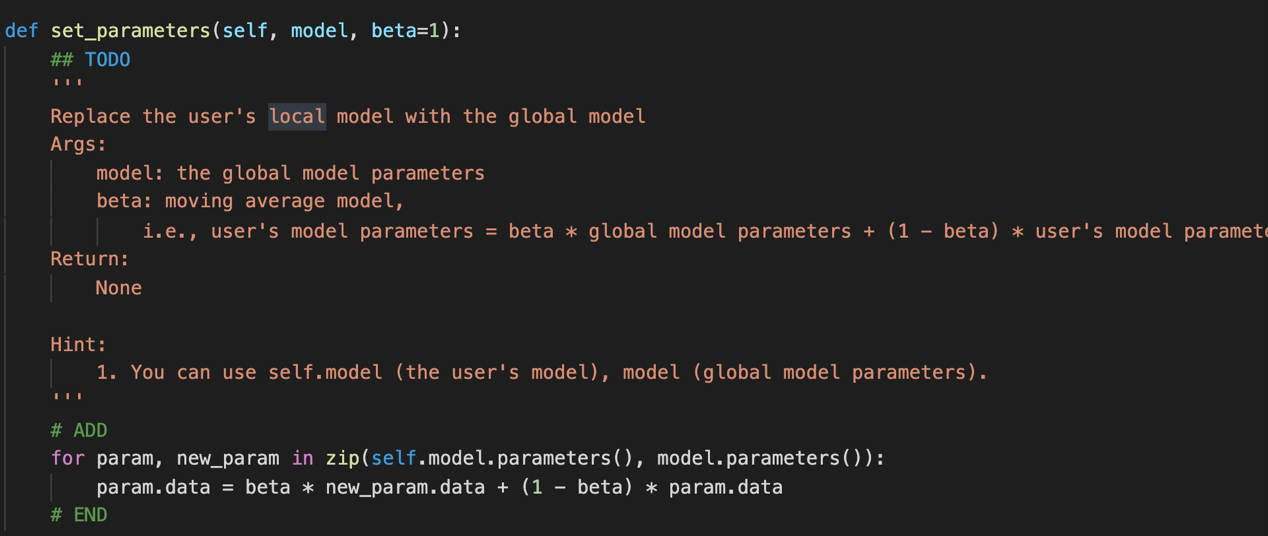
aggregate\_parameters(self)



首先先宣告一個參數列表，其形狀與model parameter的形狀一樣；接著，我們宣告total\_train\_samples去紀錄樣本總數，以利待會加權平均之計算；再來，我們計算各個user其相對應的加權平均並累加至new\_params裡頭；最後將更新後的global model parameter更新至model\_param

userbase.py

set\_parameters(self, mode, beta)

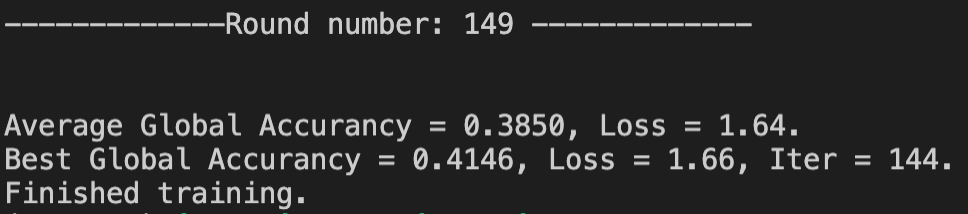


將server的model parameter送至user端的model中並使用beta控制parameter的更新方式。

1. 探討問題的原因

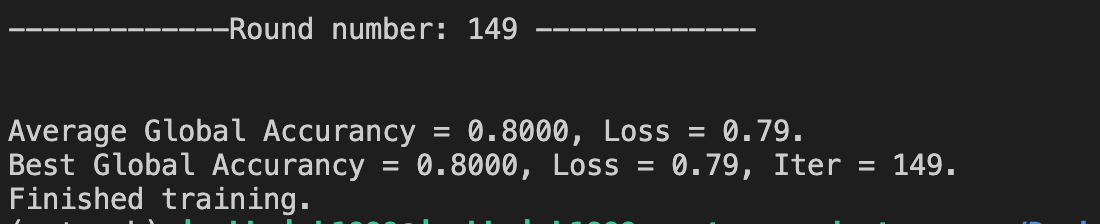
* Data Distribution

alpha 0.1



當alpha為0.1時，其數據分佈較不均勻，每個user端的data distribution更傾向於某個類別，使得在global model在不同類別上的training data不足，進而導致正確率低落。

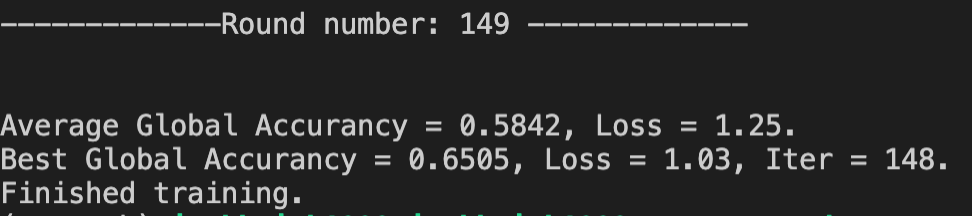
alpha 50.0



當alpha為50.0時，其數據分佈較alpha為0.1時均勻，每個user端的data distribution更加貼近整體數據分佈，在訓練global model有更好的效果，使準確率大為提升。

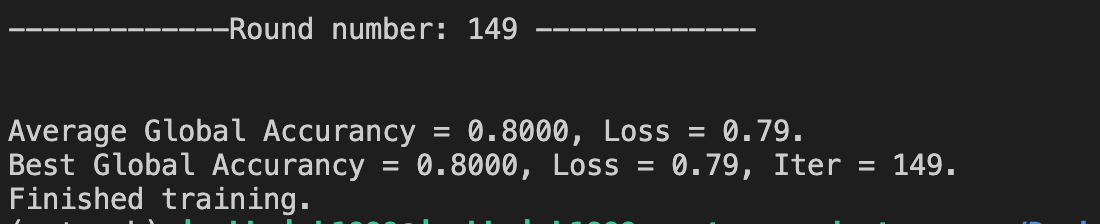
* Number of users in a round

num\_users 2：



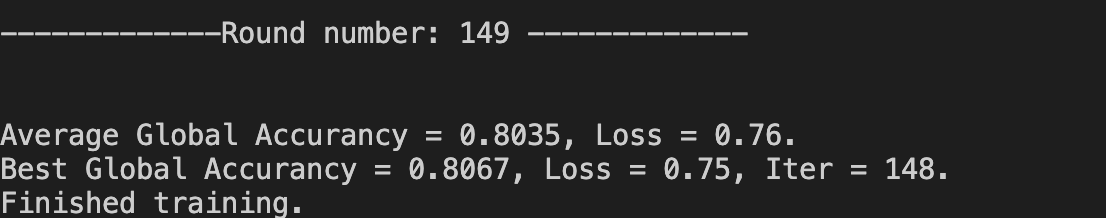
每輪訓練的user端數量較少，model parameters更新不夠充分，導致收斂速度較慢，最終準確率較低

num\_users 10：



每輪訓練的user端數量較多，model parameters更新較num\_users為2時充分，收斂速度較快，最終準確率較高

1. 最終acc的輸出截圖（--num\_users 10, --alpha 100.0）



1. 此次作業習得重點

在這次的作業中，讓我更熟悉Horizontal Federated Learning的運作原理，並且在實作select\_users, aggregate\_parameters, set\_parameters時，更了解這些function與各項參數之間是如何互相影響著global model的準確率與訓練效果。