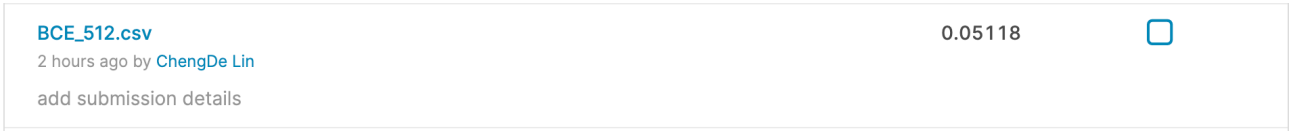
WM programming assignment 2　Report

R08922A04 林承德

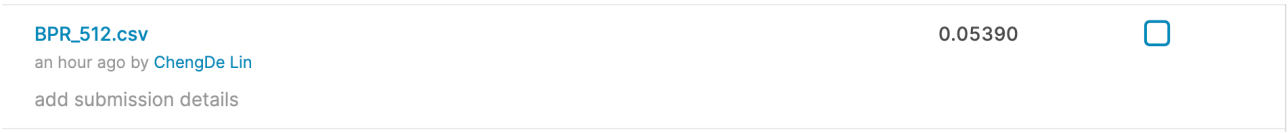
Q 1. Describe my MF with BCE

這邊我的negative sample method是隨機的從negative item中選出跟positive item的數量同的negative item，確保資料的balance。每筆sample都是triple的形式，包含有（user_id, item_id, label）此外參數的部分，這邊我使用的embedding dimension為512維（實驗發現，在32,64,128,256,512維中，512表現最好），learning rate是5e-4，同時optimizer為Adam，loss的部分是直接使用pytorch 內建的BCEWithLogitsLoss()
最後，上傳kaggle後得到的MAP為：



Q 2.

這邊我的negative sample method 是隨機的從negative item中取出與positive item數量一樣多的negative item，並將user, positive item, negative pair組合成一個triple(也可以調整取出negative item 數量，藉此增加pair的數量)此外參數的部分，和BCE一樣，經過實驗我這邊選用的是512維的embedding dimension，learning rate 為5e-4，optimizer為Adam。loss 的部分則是根據投影片上的公式來做計算。
最後，上傳至kaggle後得到的MAP為：



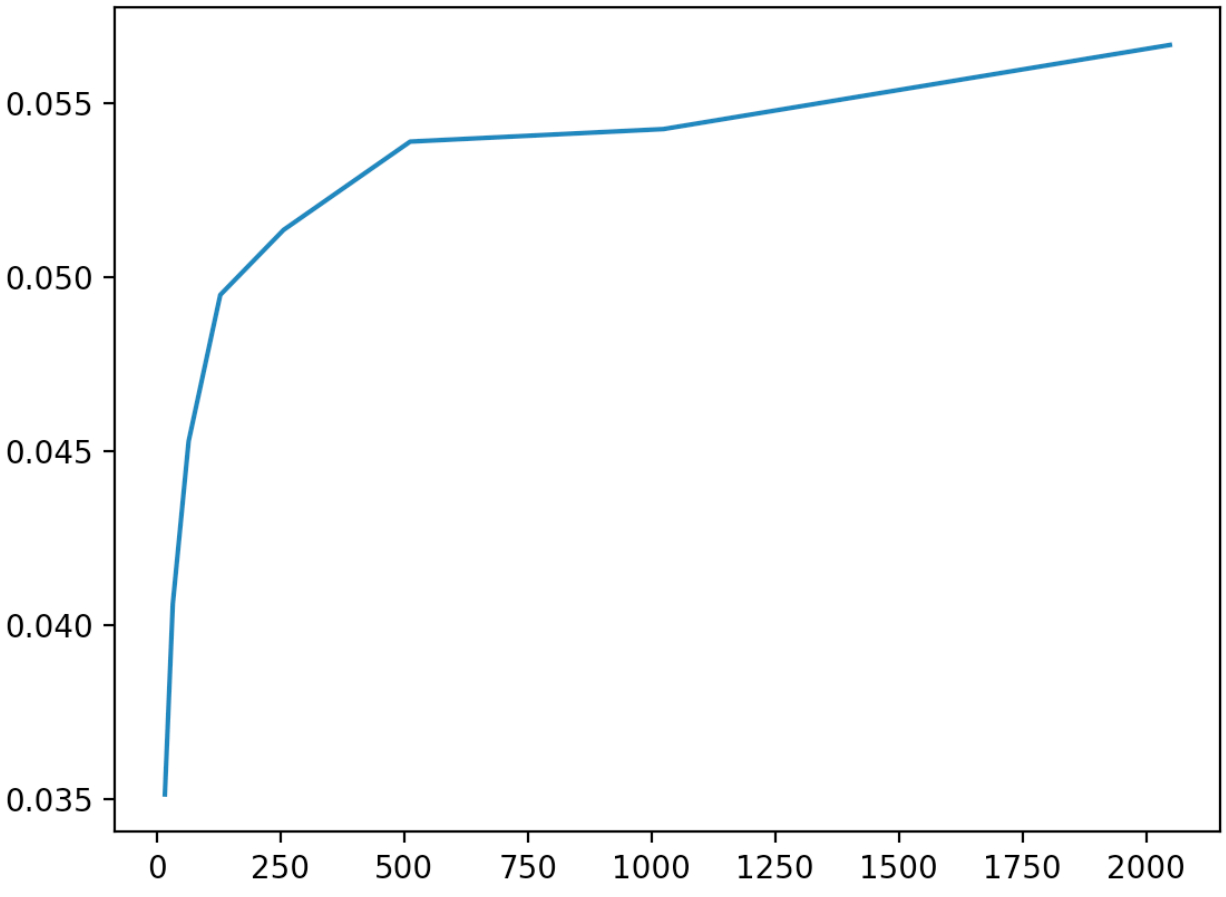
Q 3.Compare your results of Q1 and Q2. Do you think the BPR loss benefits the performance?and Why?

是的，BPR loss確實對提升performance帶來了不少的效果，我想可能是因為兩種的loss所追求的目標不太一樣所導致的。我們可以想像，BPR想做的是找出和user有著最高的適切度的item來推薦給使用者，而ＭＦ中我們所用來計算適切程度的方法就是將兩者的embedding做內積。而現在兩個loss最大的差別是，BCE想要使positive item跟negative item與user相乘的結果分別趨近1和0，而BPR則是希望這個結果分得越開越好。而這樣不同的目標可能導致，BCE透過scaling的方式，來改變embedding上的值，藉此分開positive與negative，倘若我們換用cos的方式去看則可能會發現，他其實並沒有成功的分開positive與negative。因此

當今天BPR的訴求只是要讓positive 與 negative相距越遠越好的時候，模型可能比起BCE loss更可能傾向在不scaling下，直接將兩個在向量空間分開，而這更像是我們所希望的。

Q 4.Plot the MAP curve on testing data(Kaggle) for hidden factors d = 16, 32, 64, 128

我實驗了不同的hidden factor，從d= 16,32,64.128一直到2048，其map curve 如下圖
x = [16,32, 64, 128, 256, 512, 1024, 2048]
y = [0.03512 , 0.04060, 0.04528, 0.04949, 0.05136, 0.05390, 0.05426, 0.05668]

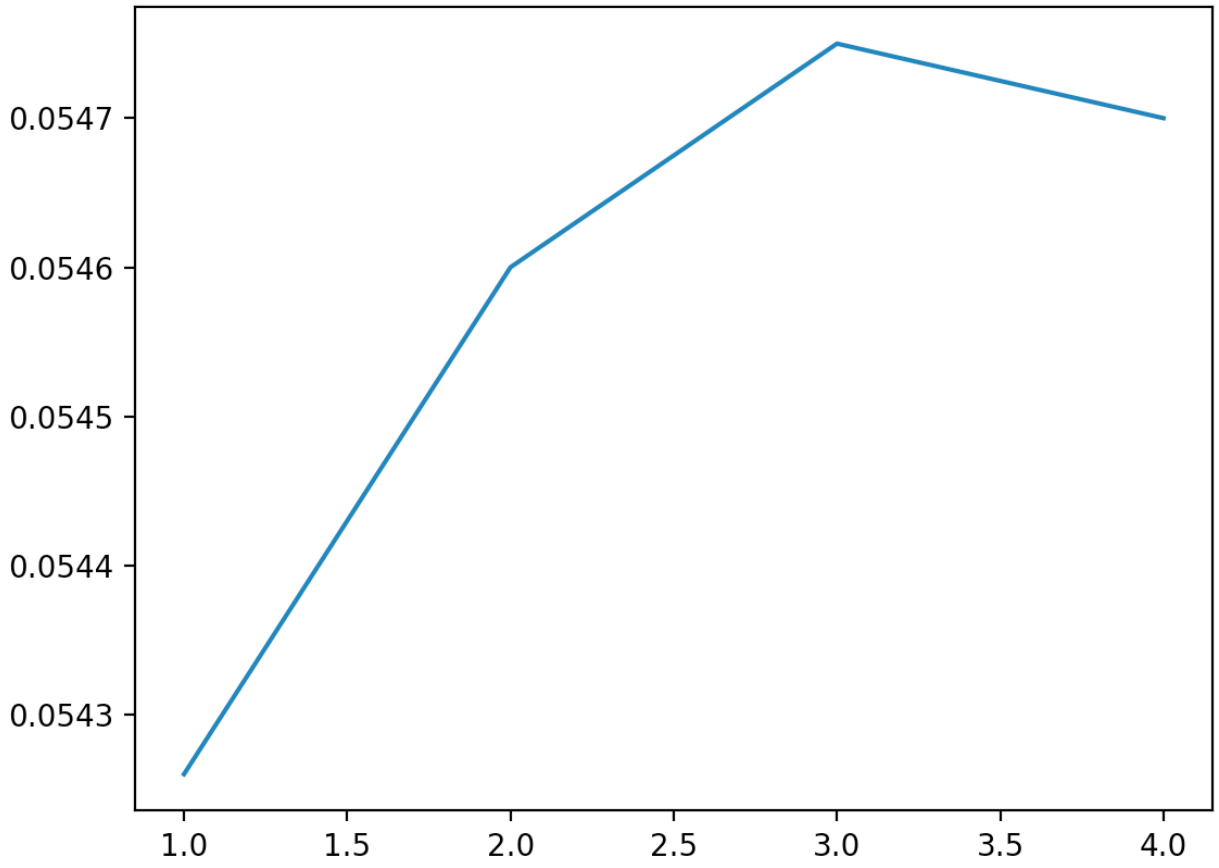


從這邊我們可以發現，隨著hidden factor的上升，map有越來越好的趨勢，尤其在16~128之前，上升的幅度極大，可能表示著較小的d不足以抓取到足夠多的資訊，來使得模型能推薦出正確的東西，而當d很大後，map上升的幅度趨緩，表示後來多出來的維度所能抓取到的更多的資訊已經漸漸的不那麼重要，所以只能及小量的提升MAP。

Q 5.Change the ratio between positive and negative pairs, compare the results and discuss your finding.

這邊我試著嘗試了一些不同的negative sample比例，從和1:1到1:4，實驗結果如下，

x = [1,2,3,4]
y = [0.05426, 0.05460, 0.05475, 0.05470]



透過調整negative sampling的比例，可以在某種程度上增加我們的訓練資料，因此對於performance的提升，是有一些幫助的。但超過一定的數量後，效果就不明顯，甚至可能導致map下降。