HW4 Discussion

資訊108 F74046496 郭旻學

使用資料為:dota2

1. 讀取資料和去除ID這項不會用到的資訊。
2. 為了減少noise，去除較無參考價值的少量資料，採用最多人玩的game mode。
3. 為了更進一步，也捨去較少人玩的game type，接著採用最多人玩的game type。
4. 完成data preparation後，以Logistic Regression製作模型與test data比對，得到整體的準確率0.5939382164367593。
5. 若將範圍集中在最多人玩的game mode 與 game type，則準確率上升至0.5954805856142584。