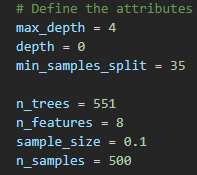
**Attributes setting**

****

Number of Trees: 551

當樹的數量成長到一定的值的時候，浮蕩的結果會趨於穩定，故我從原本幾十棵樹調整成目前551棵並train了好幾個小時得到最終結果

Number of Features: 8

根據樹的最深深度4，取了兩倍的數當做feature

Number of instance: 500

因為tree數量很多，故instance取小一點讓random的多樣性更多

**Difficulty：**

* Panda函式庫與DataFrame的不熟悉：
* 對Decision Tree的觀念不清晰
* Hyperparameters的取值
* Training過程蠻長且不能保證結果好壞

**Summarize：**

一開始在實作basic的時候，因為此次作業多使用了panda函式庫的DataFrame，導致在操作上遇到了不小的困難，它的使用方式與numpy沒有那麼相似，所以在查找文件和實際測試裡來回反復上花了不少時間，但在實作的過程中也逐漸掌握了該如何使用。

隨後在basic part裡，雖然有template的指示該往哪個方向去做，但初次接觸仍有很多觀念和想法不清晰，在這部分也是搭配著講義和網上資料將自己不足的知識都一一補全，再將這些想法逐個實現在code裡

就這麼跌跌撞撞總算完成了所有的baseline，而後就是漫無止境的training過程，因為使用random forest，每次train出來的結果有好有壞，隨後再調整參數再做training使得花費了很多時間。在多次嘗試不同的hyperparameters之後，我發現當tree的數目成長到一定數量的時候，其train出來的結果會逐漸穩定，而tree的深度和所使用的的feature有需要有一定的關係數量才可以達到比較好的結果。

完成了此次decision tree的作業，基本的machine learning的觀念想法已經大致掌握，接下來便逐漸要踏入neural network的知識領域，難度會逐漸提升，但我相信我會做好準備的。