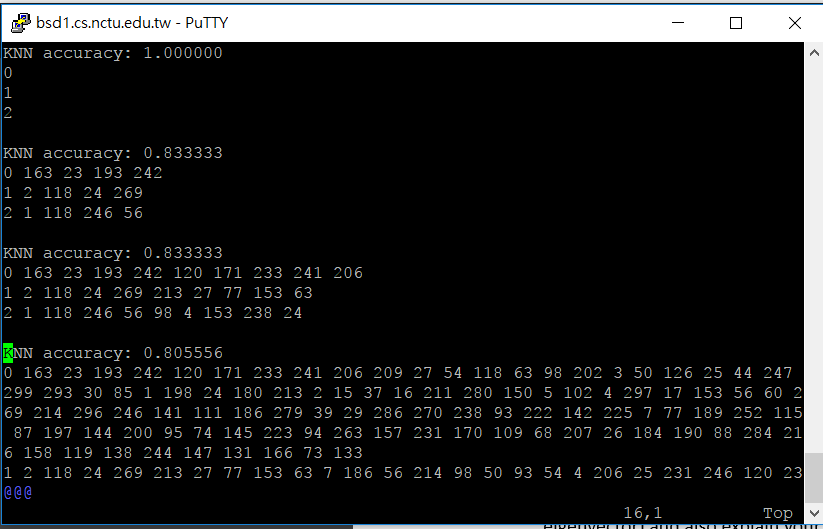
機器學習作業2 報告

0416094 黃兆宇

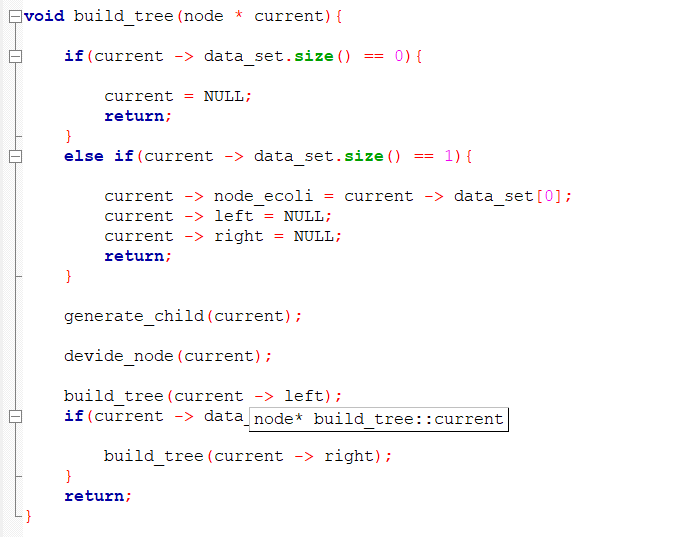
執行結果 : 

1.決定使用 k = 10 ，因為 k=1因為test case 的選擇不同，若找到極端的test case (如上圖) 會造成準確度過高，故選擇較中間準確度的 k = 10 或 k = 5 ，

最後選擇 k = 10 ，因為 k = 10 有較多參考點，結果應較 k = 5準確。

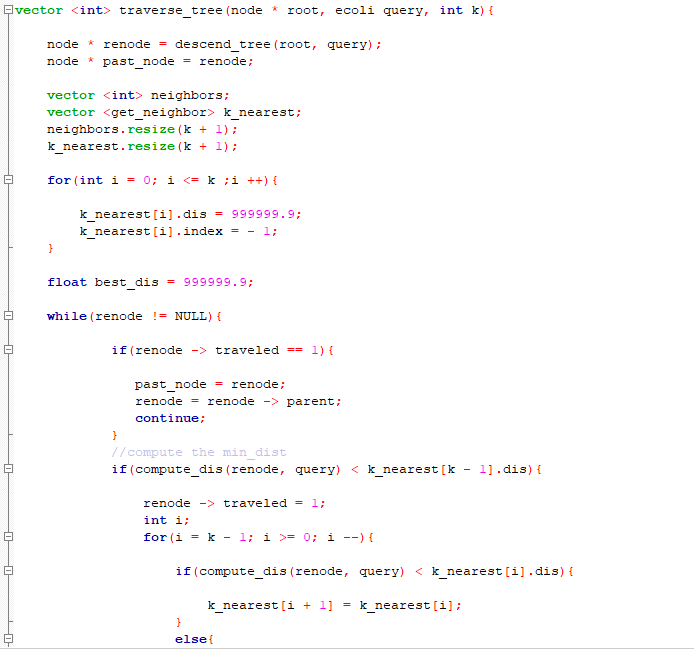
2.程式語言 : C++

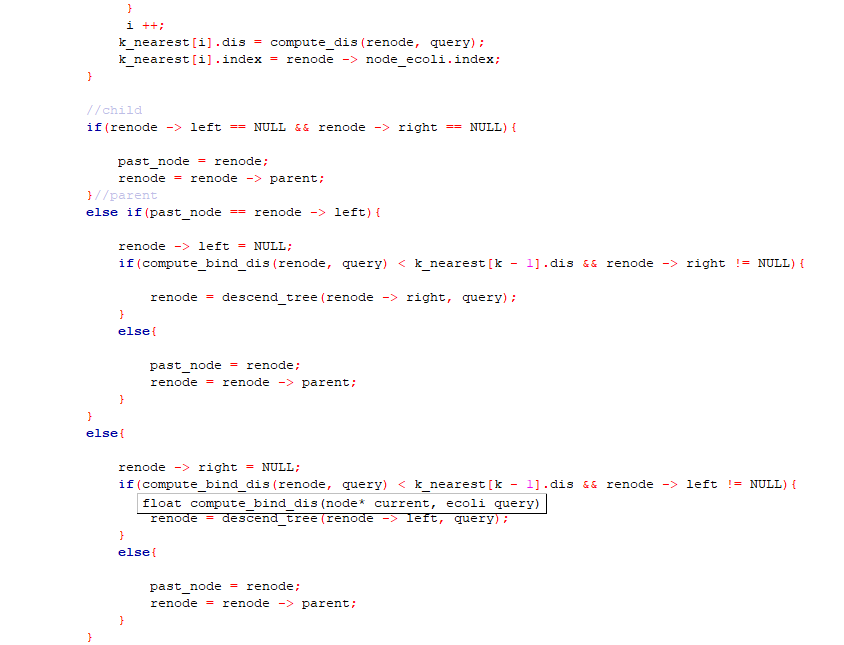
3.程式 : (function 功能說明)



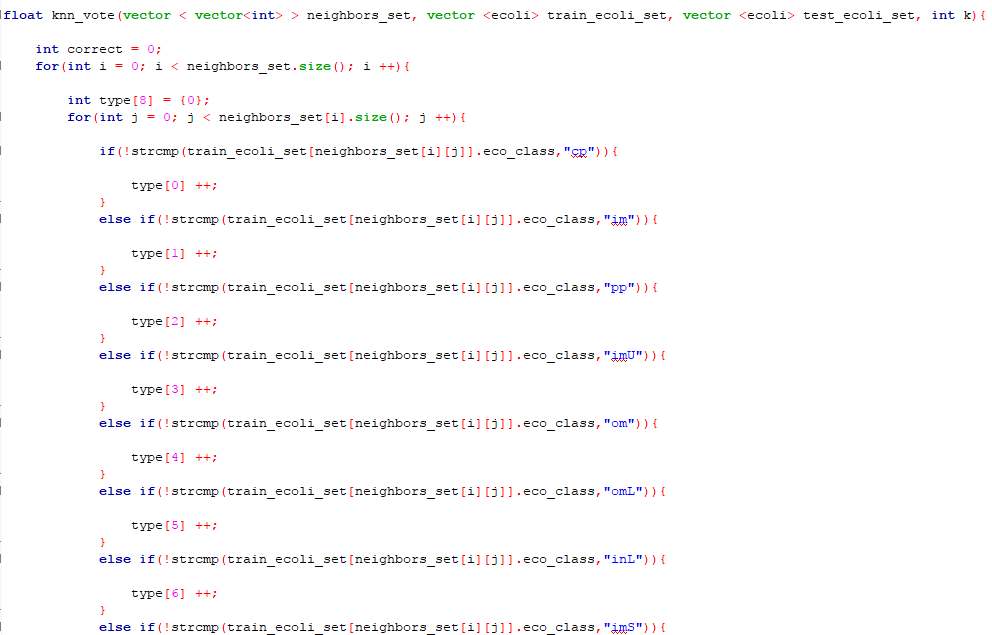
build\_tree : 建造一棵k – d tree ，主要使用建造2元樹的方式，以各feature的中位數做child 分割。

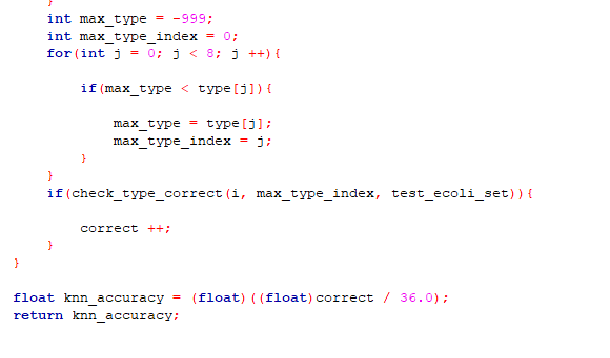
其中還呼叫gernerate\_child 和 devide\_node 兩個function，gernerate\_child是針對樹的末端對一些特例child 的特別處理(一個child)，devide\_node是分割node。





Traverse\_tree : 先以 2 分搜尋法走到leaf ， 在用back track 的方式往回走，這裡使用了k\_nearest 的 vector來儲存最接近的k個neighbors，比較特別的部分是，當query從一個child node 走至parent node 時，回將其過來方向之child pointer 指向null，也就是進行’減枝’的動作，如此能避免重複走到node，減少搜尋時間，但代價是必須重複建樹。





Knn\_vote : 最後是進行投票，將鄰近n個判斷出最多的type是為此knn樹最後的決定type，並計算accuracy。

Test\_kd\_tree : 處理test case 的輸入並將結果輸出至檔案。