

(1) random forest attributes settings

總共有 3 棵 tree

每棵樹用 5 個 feature, max_depth 皆為 6

第 1 棵用了 2500 個 instance

第 2 棵用了 2000 個 instance

第 3 棵用了 1500 個 instance

(2)遇到的困難

當初在設計 `make_partition()`的回傳值時，只回傳含 2 個 column(feature, "diabetes")的 dataframe，一開始不知道為什麼，但後來發現會在 `build tree()`呼叫到 `find best split()`的時候(我在 `find best split` 會 loop through 所有的 attribute)產生 `KeyError`

(3)解決方法

上面的 bug 我 trace 了一下才發現這個問題，最後決定 `find best split()`回傳的 dataframe 保留所有的 attribute 就不會有這個問題。一開始寫的時候沒料到會發生這個問題，有些 Error 是因為沒注意到某 function 的 return value 傳到下一個 function 處理時會不會發生問題所導致，這是在 hw2 學到的細節。

(4)建議

希望以後 template 的 indentation 可以有 Tab 的版本，兩個空白實在不習慣