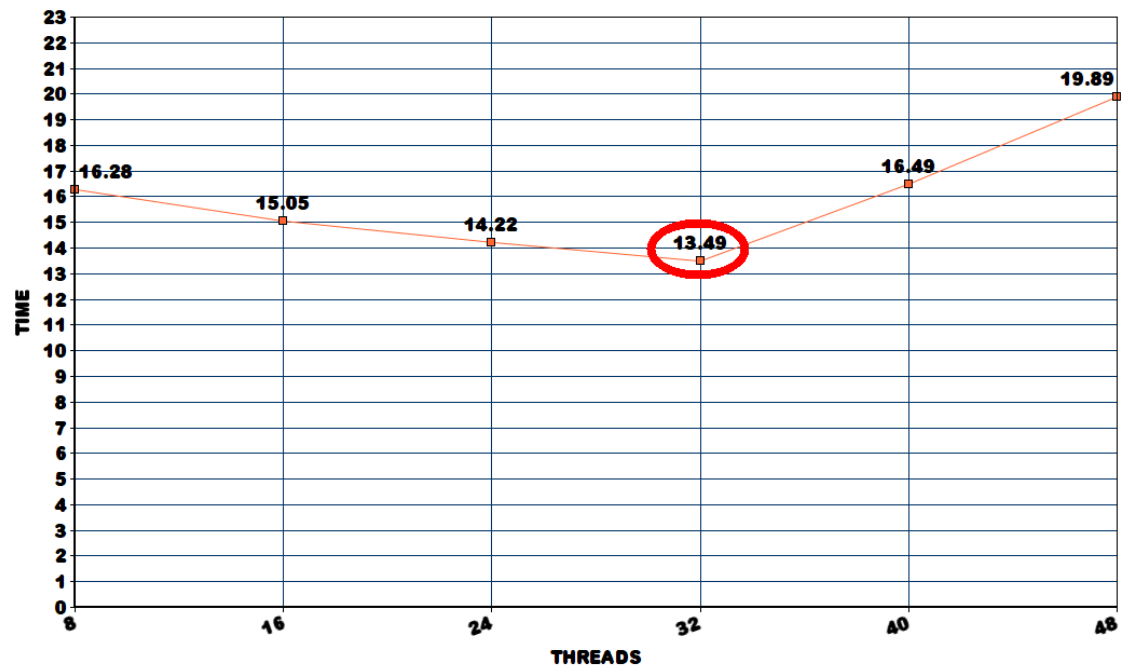


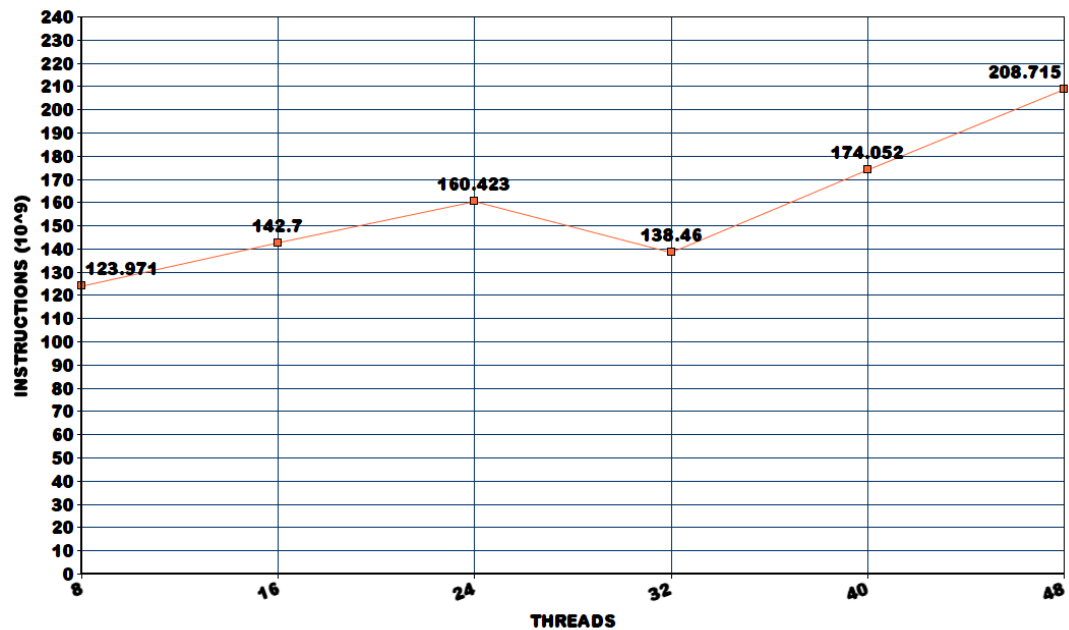
## B04902112 張凱捷 SP HW4 Report

(b) 在建 random forest 的時候可以開 multi thread，讓每個 thread 都建一些樹，當 tree size 變大時 multi thread 效果會變顯著

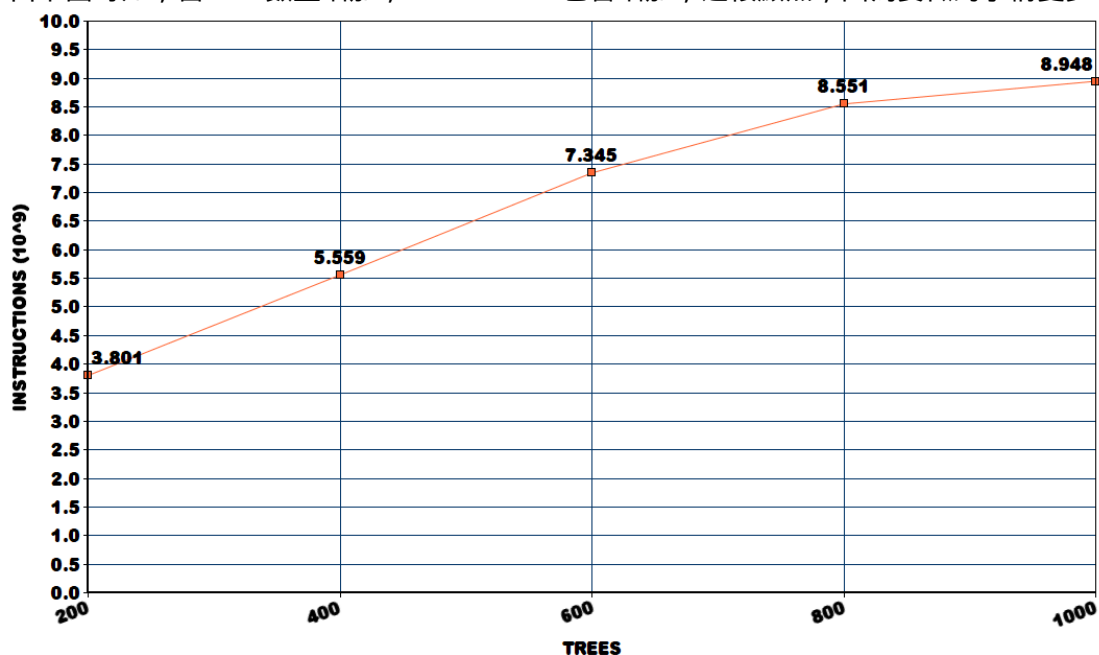
(c) 以 tree size = 512, tree number = 50 來做測試  
由下圖可以發現 thread 並不是爽爽開，開 thread 也是有 cost，因為他們之間要做 context switch，又因為你的 CPU 核心數有限，所以 thread 大於一定數量後反而會變慢



(d) 以 tree size = 512, tree number = 50 來做測試  
建樹的 instruction，加總理論上會差不多，但是多了一些沒太大功用的 thread，建 thread 要成本，所以 thread 增加時 instruction 可能會增加



- (e) 以 tree size = 33, thread number = 50 來做測試  
由下圖可知，當 tree 數量增加，instructions 也會增加，這很顯然，因為要做的事情變多了



- (f) 就我自己的觀察，tree size = 33 = feature 數量時正確最高，tree 數量增加時正確率會比較穩定，後者感覺很顯然，但是前者我一直覺得怪怪的，感覺取越多資料近來應該會愈準才對（？說不定是我程式寫爛或是 data 太爛，反正我試了一下正確率平均大概 89%，應該沒有神麼太大的 bug。