P76074282林品磊 Project2\_report

1. Dataset：
   * 輸出.csv：  
     我用爬蟲爬全台灣已上市公司的基本資料

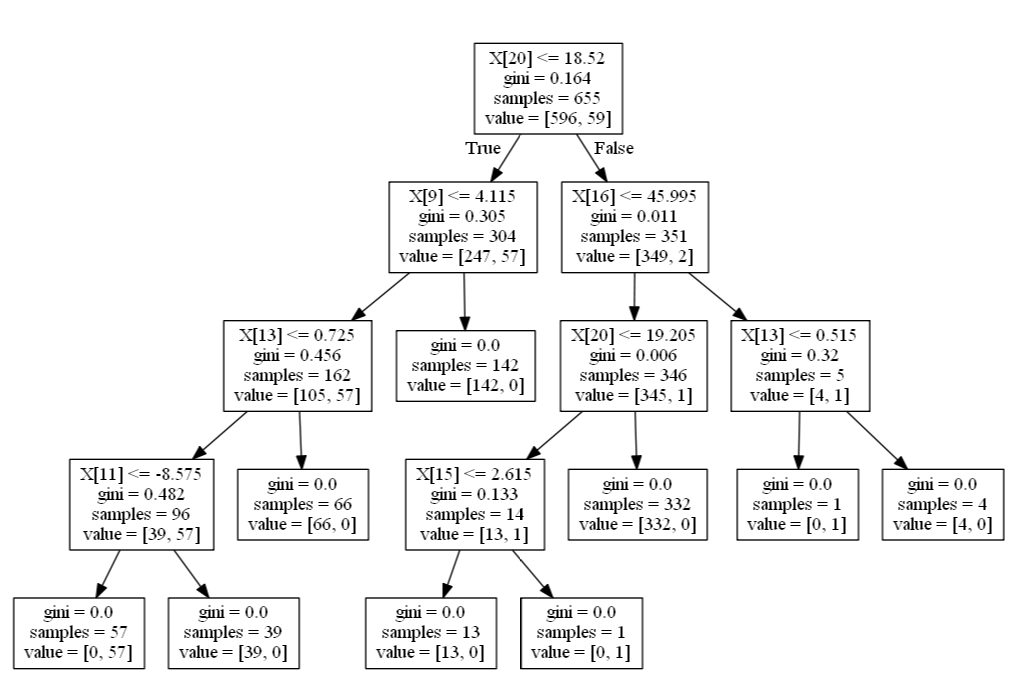
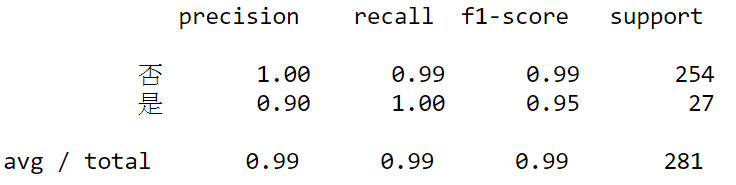
Ex. http://jsjustweb.jihsun.com.tw/z/zc/zca/zca\_1103.djhtm

* + Attribute：

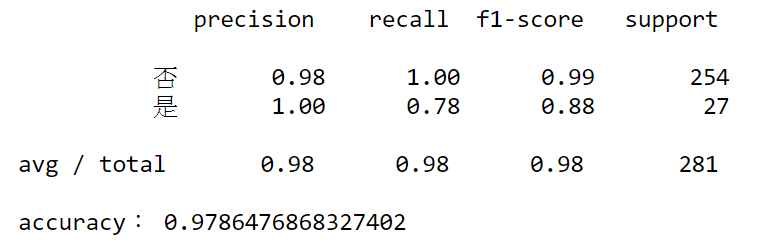
1. 公司代碼
2. 開盤價
3. 最高價
4. 最低價
5. 收盤價
6. 漲跌
7. 本益比
8. 成交量
9. 同業平均本益比
10. 殖利率
11. 總市值
12. 今年以來(投資報酬率)
13. 每股淨值(元)
14. 貝他值
15. 每人營收(仟元)
16. 標準差
17. 每股營收(元)
18. 負債比例
19. 股價淨值比
20. 股本(億, 台幣)
21. 營業毛利率
22. 現金股利(元)
23. 營業利益率
24. 稅前淨利率
25. 資產報酬率
26. 股東權益報酬率
    * Rule：
      1. 本益比<平均
      2. 貝他值<平均
      3. 殖利率>平均
      4. 今年以來(投資報酬率)>平均
      5. 營業毛利率>平均

🡺符合上述五點：可投資。否則，不可投資

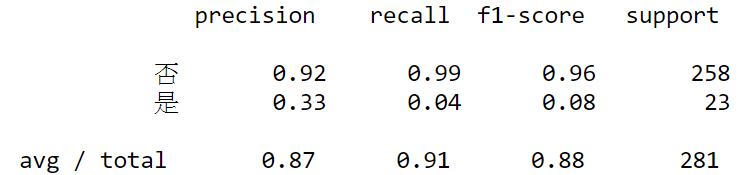
1. 結果：
   * Decision tree：  
     p.s因為中文顯示會變亂碼，所以要去對照data的欄位了><



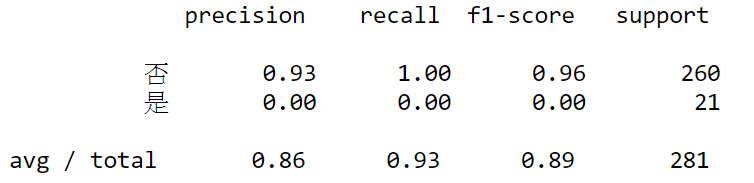
* + Random forest：



* + kNN：



* + SVM：



1. 心得：  
   現在要使用分類的model真的很方便，很多套件都已經寫好了  
   沒想到原來decision tree跟 random forest的準確度竟然如此高  
   因為不想只是為了寫作業而寫，所以我就自己爬了真實的資料  
   之後還可以真的應用在選股上☺