

學號:0411509 姓名:許家維

1. 選用 DATA

banknote authentication Data Set

- <http://archive.ics.uci.edu/ml/datasets/banknote+authentication>

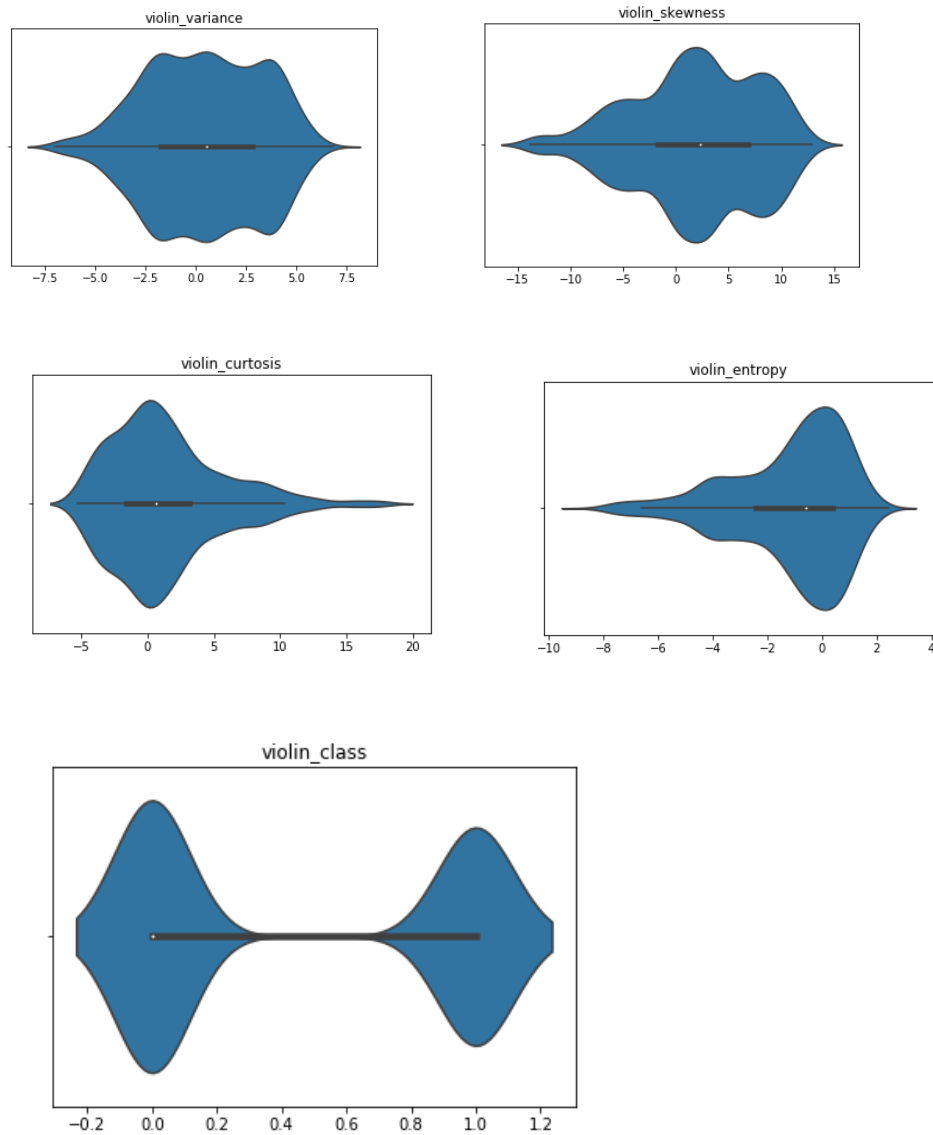
他的參數主要是透過 Wavelet Transformed 得到鈔票圖像的

- A. Variance
- B. Skewness
- C. Kurtosis

並也把圖像亂度(entropy)也當成 DATA,

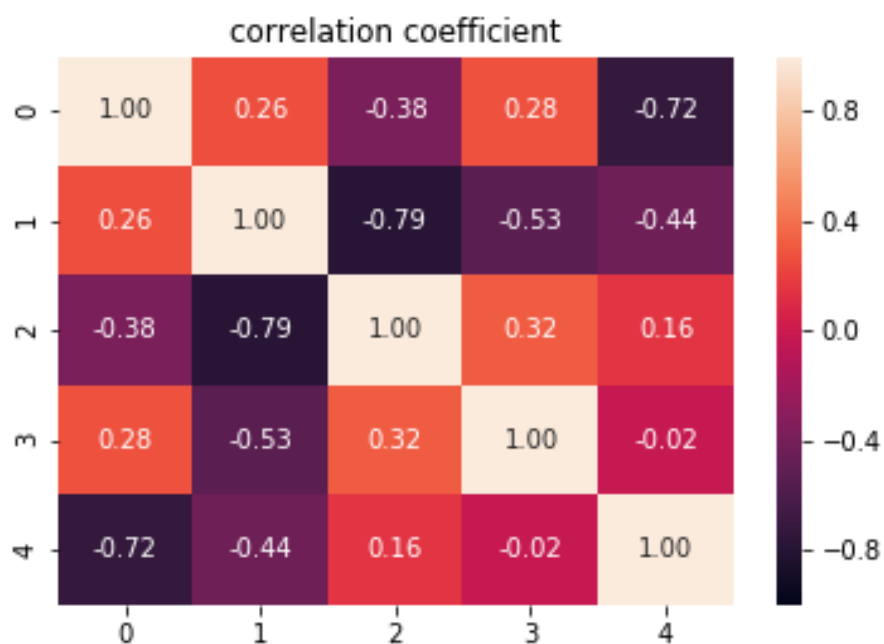
目標設定為能夠辨識出真假鈔票

2. DATA 分布狀態



數據有些有跨數量級的差距,可能需要對總體做 z-score normalization 已得到最好的特徵
數據分布也並無太奇怪的地方都大致符合統計假設

3. DATA 相關係數



所有數據項目包含分類向對其做相關係數作圖,發現數據間幾乎無高度相關的參數,代表每個參數不是完全毫無意義就是數據間都有完整的涵義不能被其他項參數所取代

4. 初步分析

利用 ISOMAP 先將數據降維,此次降成 2 維並用鈔票分類作為數據點顏色,查看資料是否可被輕易分類,實作結果如下,數據很明顯得分維 2 類

