

DM HW1

PCA and NMF

FER2013

### [Dataset]

FER2013 from Kaggle competition

<https://www.kaggle.com/c/challenges-in-representation-learning-facial-expression-recognition-challenge/data>

資料為 48 \* 48 pixel 的灰階人臉情緒圖片，總共有七總情緒  
(0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, 6=Neutral)

我們拿 Happy 和 Sad，shuffle 並各 sample 出 5000 筆來做  
(face\_sample.png)



## [Languages and Toolkit]

Python3

scikit - learn

XGBoost

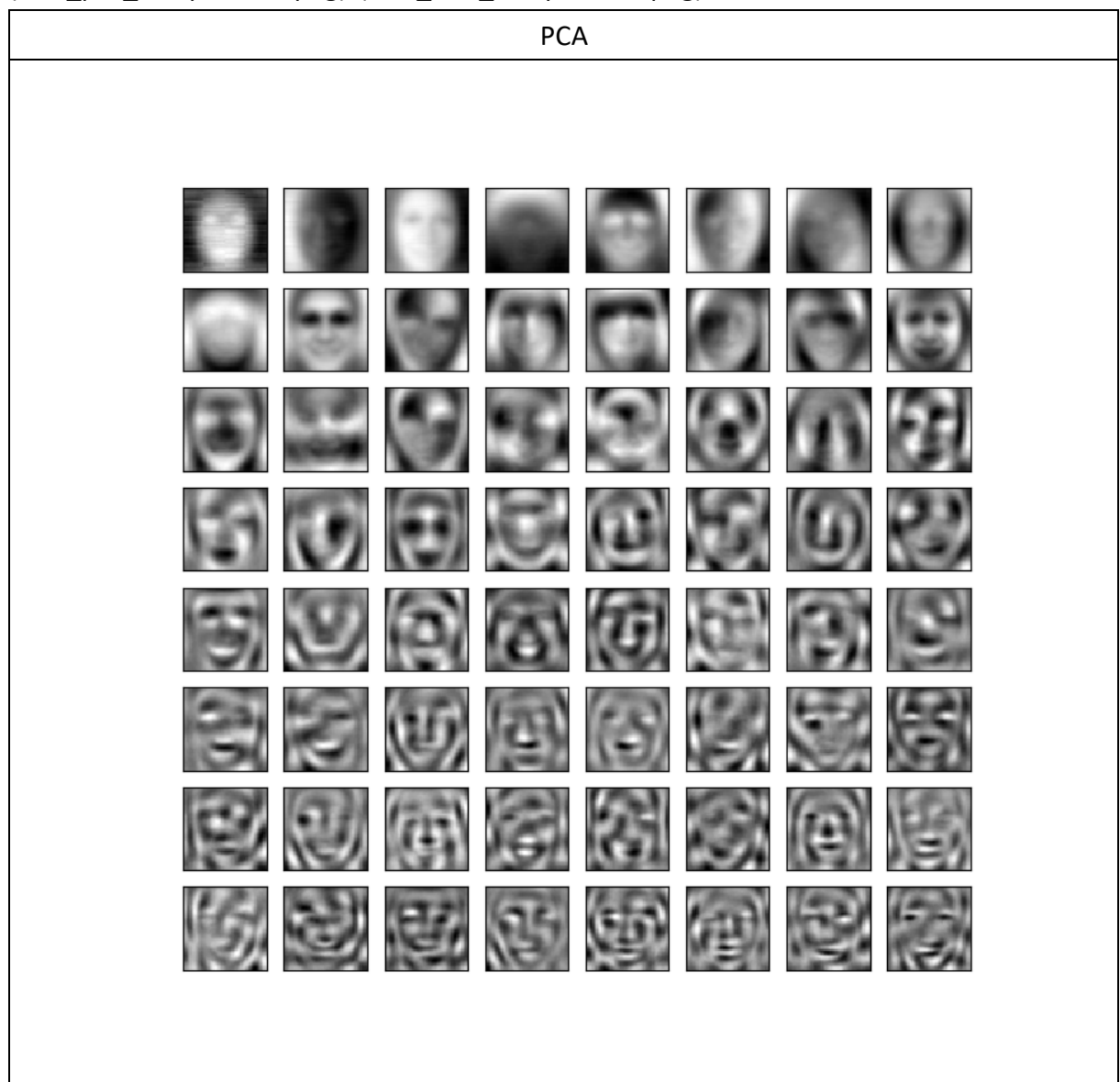
## [PCA and NMF component]

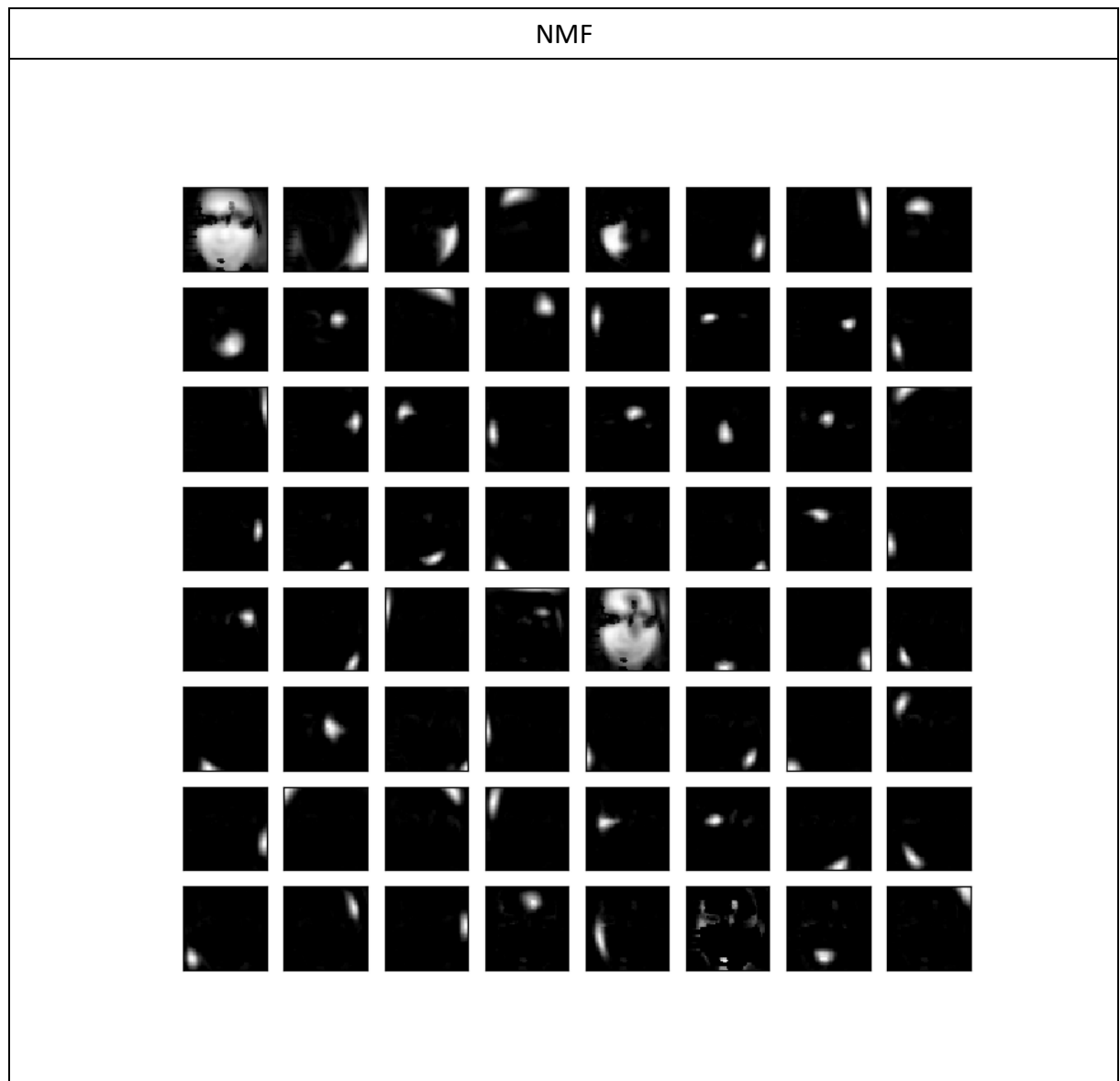
首先，我們嘗試將圖片透過 PCA 和 NMF 壓縮成 64 維度

並觀察 PCA 和 NMF 取出的 components，並做成圖

(順序由左至右、由上至下)

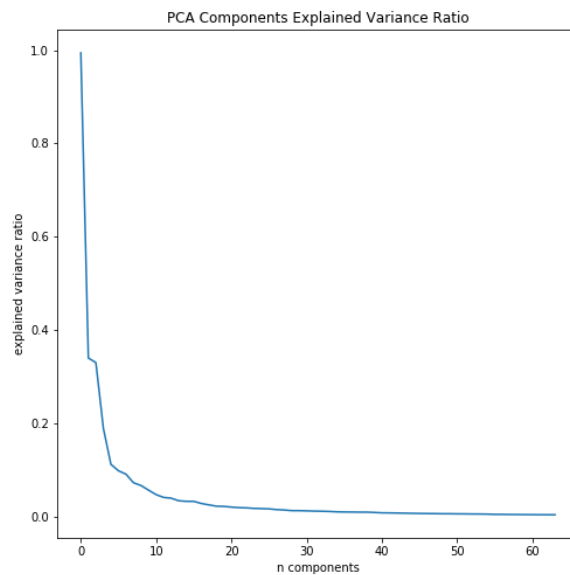
(face\_pca\_components.png)/(face\_nmf\_components.png)





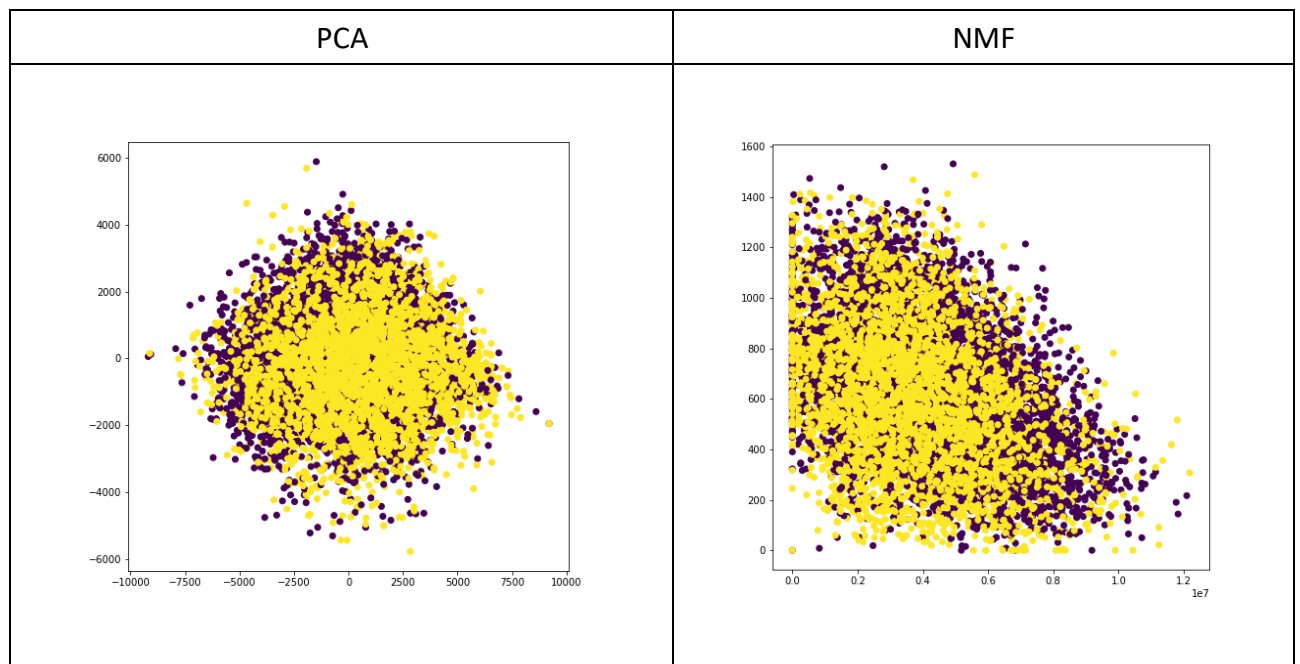
- (1) PCA 所取的成分相較於 NMF 更為人類能直觀所理解  
（圖片中人臉的輪廓較為明顯）
- (2) PCA 和 NMF 都是越上層的 components 越有人臉的輪廓

另外，我們也觀察了 PCA 解釋的變異量  
(face\_pca\_explained\_varinace\_ratio.png)



可以發現越後方的 components 能解釋的變異量越少，為指數衰減

接著我們透過 PCA 和 NMF 將資料降成二維並做 scatter plot  
(face\_pca\_scatter.png)/ (face\_nmf\_scatter.png)



降到二維後，雖然不是很明顯，但仍可發現兩種情緒在分佈會有稍微的不同。

## [PCA and NMF on Classifier Performance]

我們選用了幾個分類模型

並個別做 10-Fold 來檢驗 PCA 和 NMF 的降維效果

選用之 Classifier

SVM

RandomForest

ExtraTrees

XGBoost

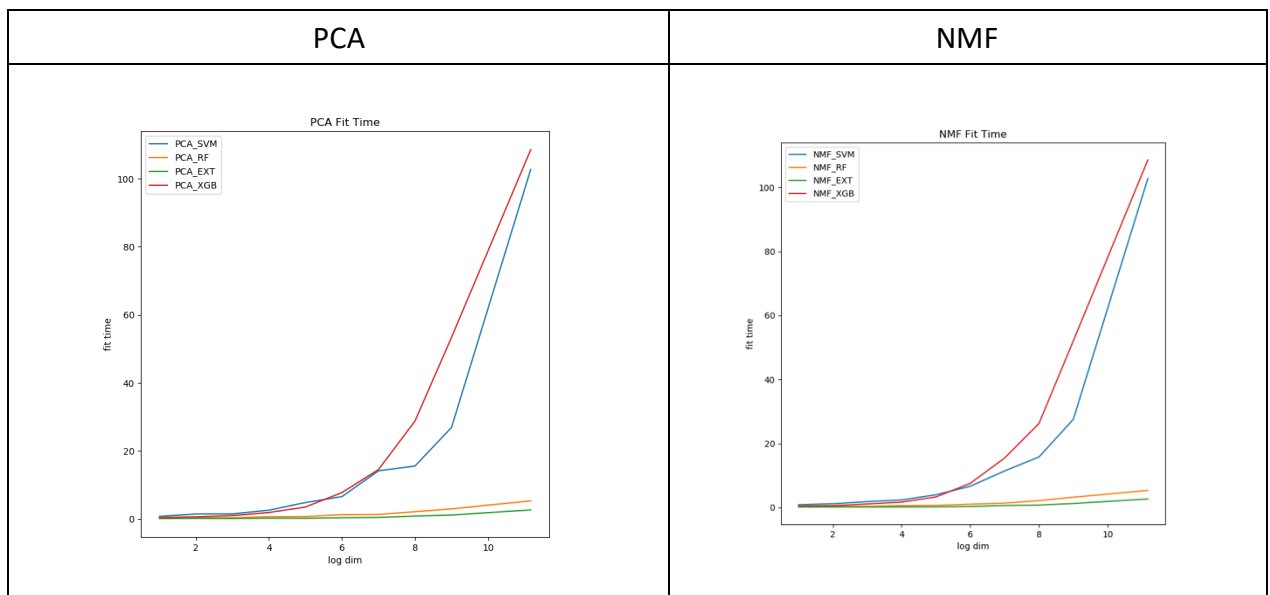
檢驗之維度

[2304, 512, 256, 128, 64, 32, 16, 8, 4, 2]

(2304 即  $48 * 48$  的原始維度，沒有經過降維)

### 1. Fit time

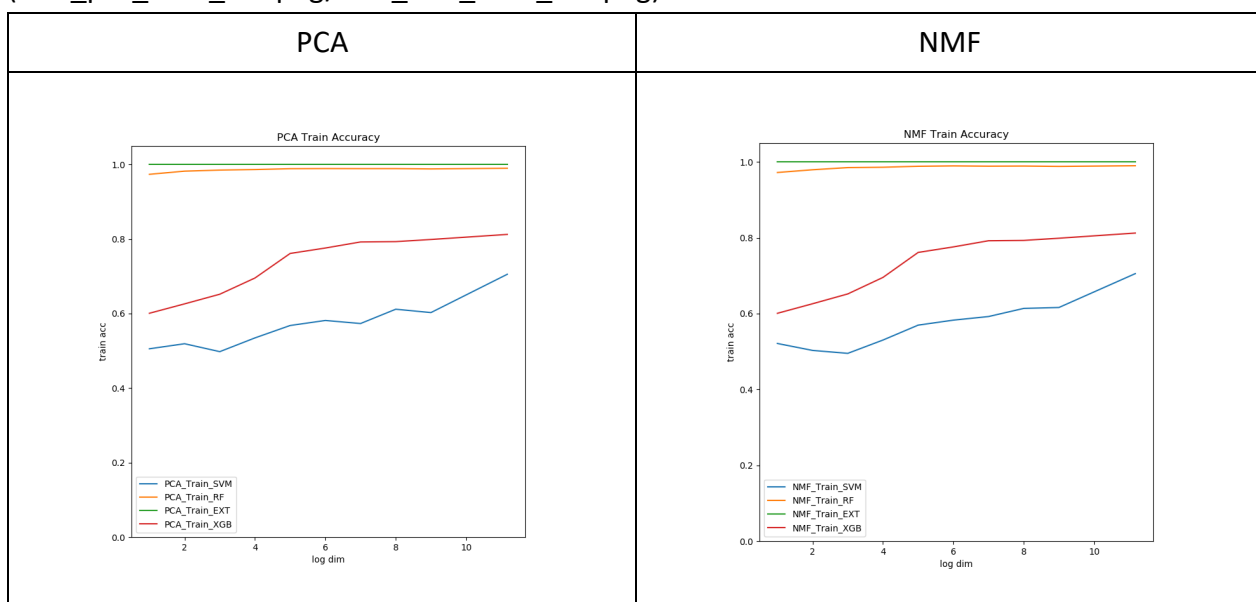
(face\_pca\_fit\_time.png/face\_nmf\_fit\_time.png)



(1) 可以發現降維後 SVM 和 XGB 大幅度減少了 model 訓練時間  
而 RF 和 EXT 也有明顯下降

## 2. Accuracy

(face\_pca\_train\_acc.png/face\_nmf\_train\_acc.png)



(1) 在 train acc 上，RF 和 EXT 都趨近於 1.0

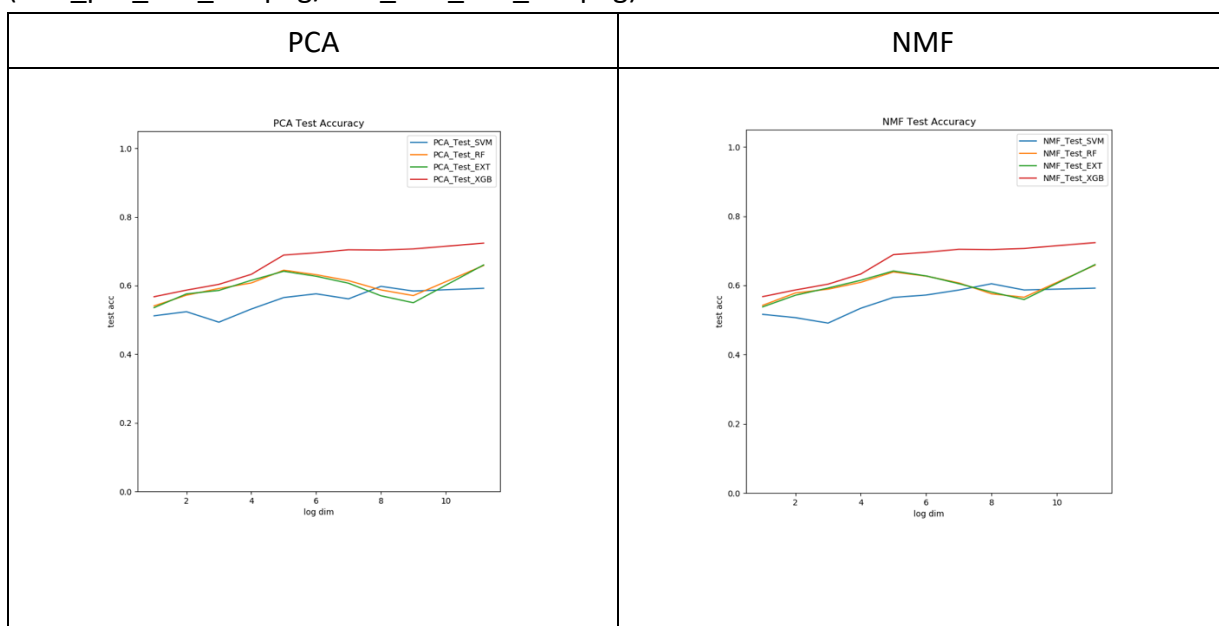
跟後面的 test acc 比較可以知道發生了 overfitting

XGB 大概可以達到 0.6~0.8 的準確率

SVM 則是稍微差了一些約 0.5~0.7

(2) 整體而言，可以發現維度降得越低，XGB 和 SVM 的 train acc 也會跟著越低，不過對於原本就 overfitting 的 RF 和 EXT 就比較沒有影響

(face\_pca\_test\_acc.png/face\_nmf\_test\_acc.png)



(1) test acc，我們可以發現到表現最好的是 XGB，約 0.6~0.7

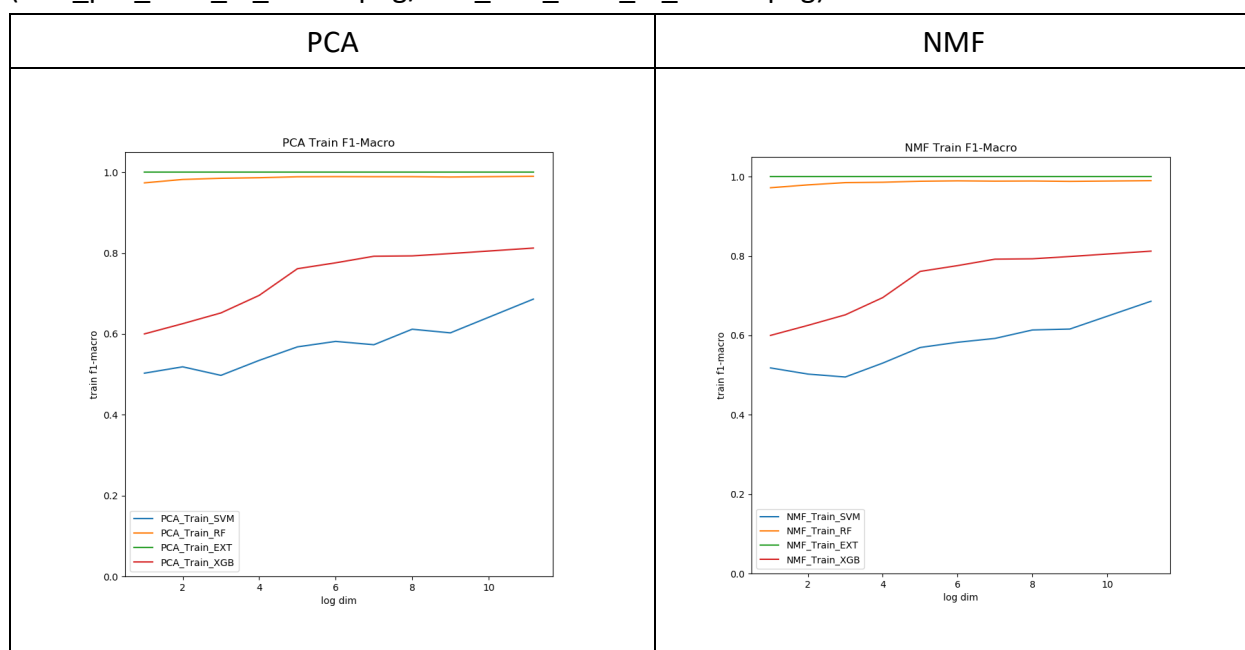
RF 和 EXT 次之，約 0.6 上下

SVM 最不好，約 0.5~0.6

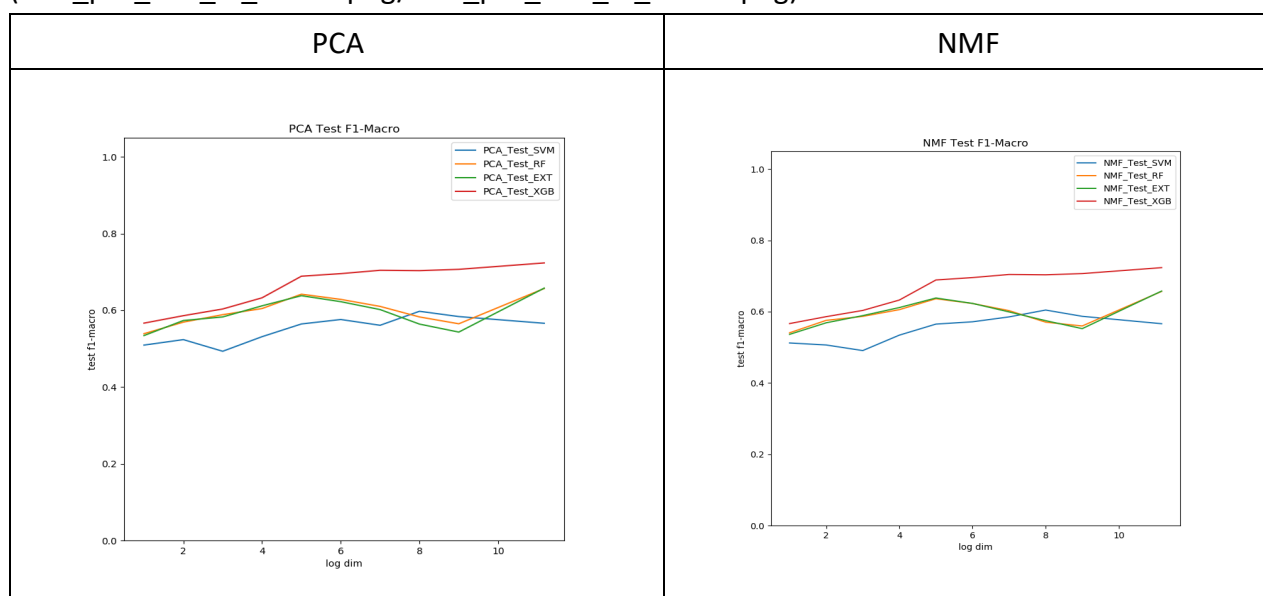
(2) 維度越低，test acc 也跟著下降

### 3. F1-Score

(face\_pca\_train\_f1\_macro.png/face\_nmf\_train\_f1\_macro.png)



(face\_pca\_test\_f1\_macro.png/face\_pca\_nmf\_f1\_macro.png)



#### 4. Decomposition Time (in seconds)

Decomposition Dim	PCA	NMF
512	13.5825	1205.2244
256	6.4271	438.1547
128	11.0945	200.6971
64	3.1308	69.4797
32	4.0512	39.9522
16	2.0576	26.4505
8	1.6477	19.1031
4	1.8589	18.4982
2	1.5582	20.4613

(1) PCA 降維速度比 NMF 快很多

(2) 基本上，降的維度愈低，所需收斂時間愈短

#### [Conclusion]

透過 PCA 和 NMF 降維，可以降低訓練時間（特別是對 XGB 和 SVM）  
而 trade-off 則是準確率也會下降。

因此，在做機器學習時，如果今天的 feature 數量非常多，（如 image pixels、text tfidf、超商中顧客對產品的購買資訊 ... 等等）我們就可以考慮使用 PCA 或 NMF 來進行降維，並在盡量不失去準確度的情況下，選擇最合適的維度來減少 model 所需的 training 時間。