110590007 資工系大三 白宸安

1. problems you encountered:

遇上的主要是在實作MY\_PCA()和MY\_SparsePCA()這兩個class，因為有較多的數學公式，要將它們理解並轉換成code，花了不少時間。另外在implement advanced part時嘗試了不同的component數量，發現較大的值會保留更多資訊，能使f1\_score更好。

1. how you implement the basic PCA
2. x\_train\_cent = x\_train - mean\_train來做centeralize data。
3. 把cent data用np.dot來做covariance matrix。
4. 用np.linalg.eigh(cov\_mat)來算eigen values和eigen vector
5. 用np.argsort(eigen\_val)[::-1]來sort eigen values
6. 選擇前n個為component：self.eigen\_vec[:, :self.n\_components]
7. Transform the data：data\_pca = np.dot(centerized\_data, self.principal\_components)
8. Plot eigenvectors and the reconstruct image

Reconstruct image: 一張含有 螢幕擷取畫面, 圓球 的圖片

自動產生的描述eigen vectors: 一張含有 螢幕擷取畫面, 圓球 的圖片

自動產生的描述一張含有 螢幕擷取畫面, 圓球 的圖片

自動產生的描述

1. Plot the distribution of eigenvalues

一張含有 文字, 螢幕擷取畫面, 行, 圖表 的圖片

自動產生的描述

1. Introduce the preprocessing/PCA methods you implemented in advanced part.

在advanced part裡面，我調高了component的數量，調成500，讓資訊能被保留的更多，這樣也能從dataset裡學到更多資訊，有發現增加component數量能使f1\_score變高，但大約有個閥值，超過之後的component增加就沒有什麼效益了。