

Analysis report

Pak Wah Wong

Centennial College

Assignment #1

Question 1

2. Display each digit

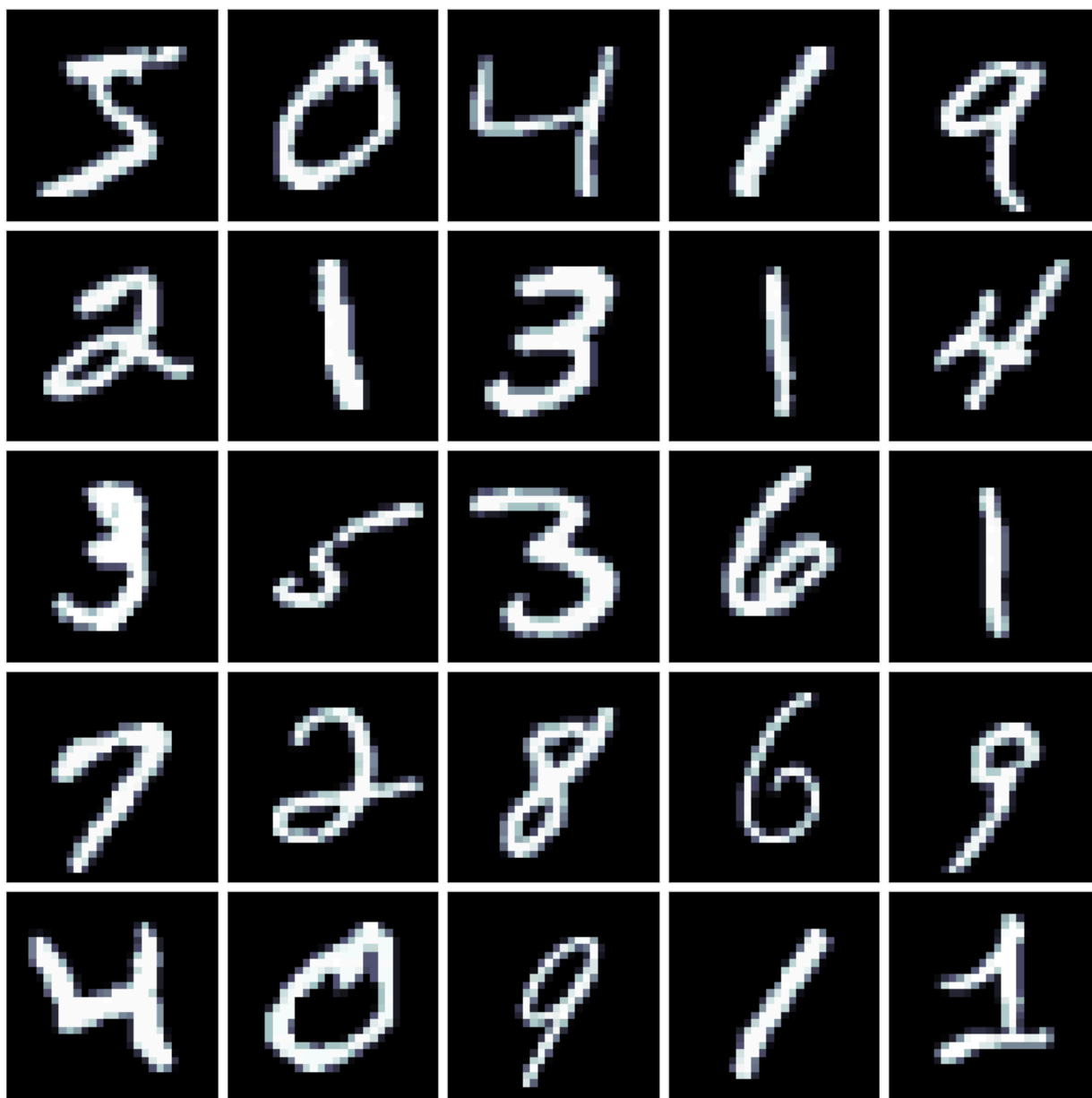


Fig. 1 First 25 digit of mnist_784

3. Explained variance ratio of 1st principal component: 0.09746116

Explained variance ratio of 2nd principal component: 0.07155445

Since the sum of explained variance ratio for the first two components is only 0.16. It is insufficient to preserve the variance.

Assignment #1

4.

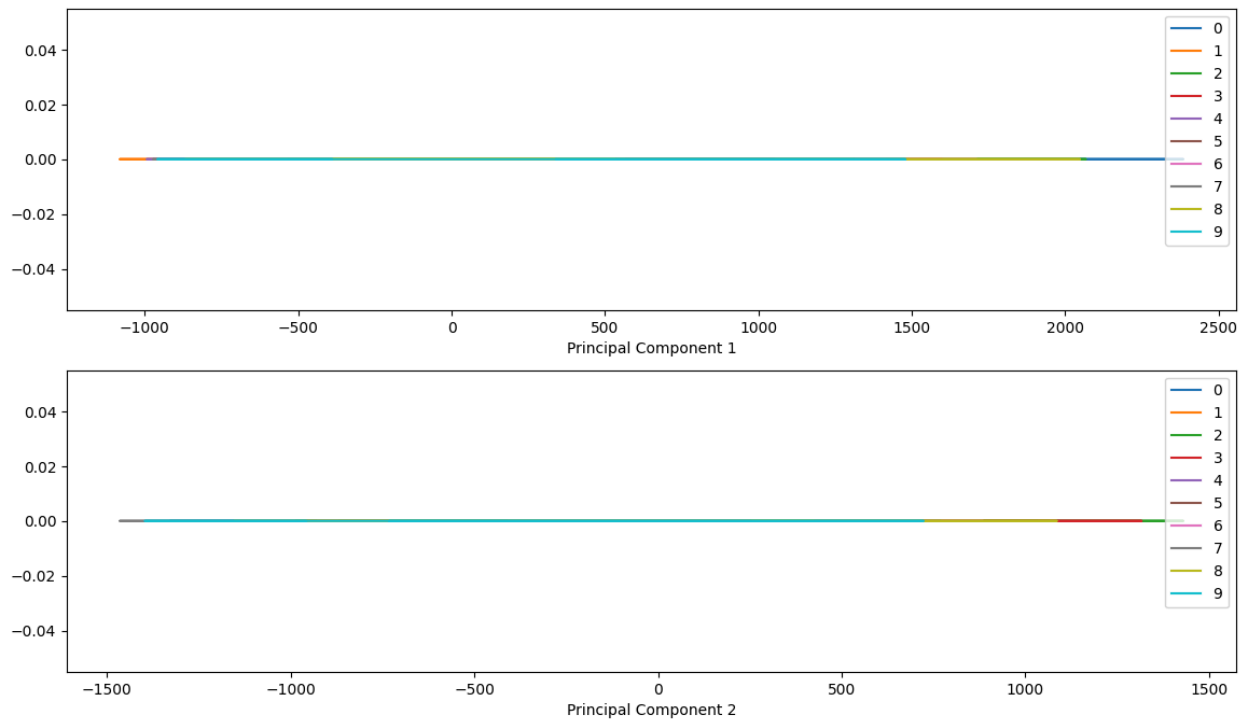


Fig 2. 1D hyperplane of 1st and 2nd principal component.

6. Display each digit. Digit after compress has less details and noise but is still easy to recognize.



Fig3. 1st row is the original digit. 2nd row is the compressed digit using PCA with 154 dimensions.

Assignment #1

Question 2

2.

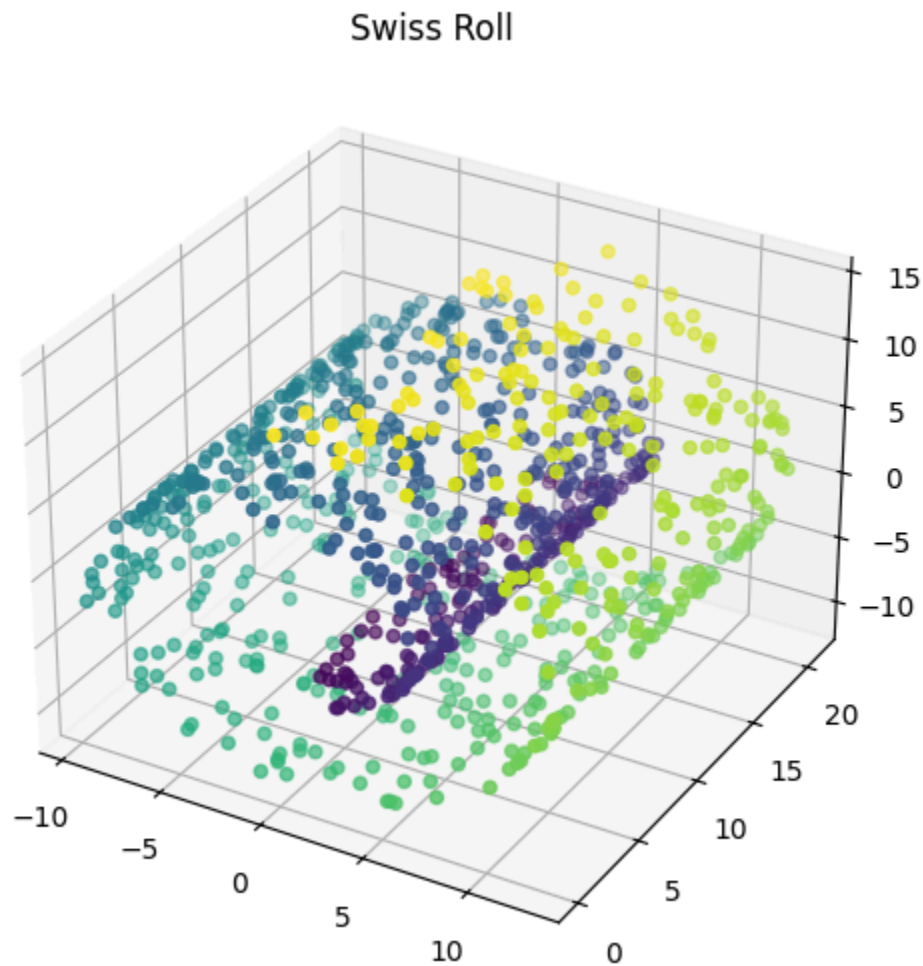


Fig 4. Scatter plot of Swiss roll dataset

4.

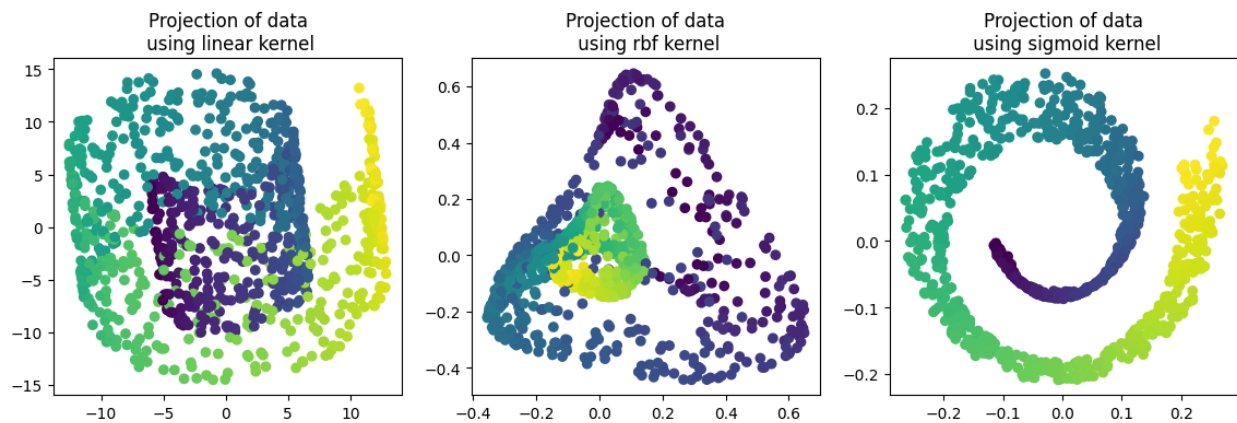


Fig 5. Projection using linear, RBF and sigmoid kernel

Assignment #1

5. Best gamma and kernel found by GridSearchCV: {'kpca__gamma': np.float64(0.03), 'kpca__kernel': 'sigmoid'}

6.

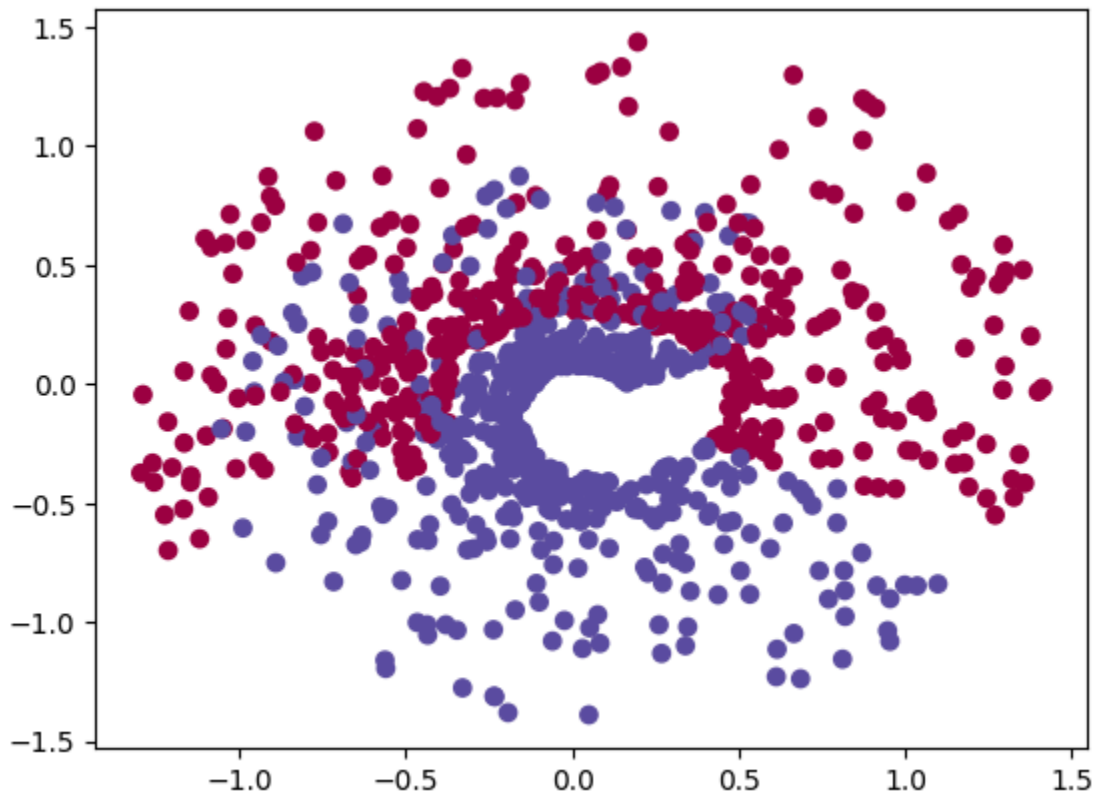


Fig 6. Scatter plot from the result of GridSearch CV