Project 1. Face Recognition

Name: Chang-Hong Chen

UID: 117397857

Neutral Face vs. Face Expression

1. With original data

	Train accuracy	Test accuracy
Bayesian Classifier	0.8843	0.75
KNN (k=3)	0.8469	0.8125
RBF kernel (σ=3)	0.8875	0.8875
Polynomial kernel (γ=2)	0.62875	0.6125

- KNN k tuning

	Train accuracy	Test accuracy
k=1	1	0.7625
k=3	0.8469	0.8125

- RBF kernel σ tuning

  RBF kernel SVM is optimized by gradient descent.
  This tuning is examined by validation data.

	Validation accuracy
σ=3	0.59375
σ=5	0.5
σ=7	0.5

- Polynomial kernel γ tuning

  Polynomial kernel SVM is optimized by gradient descent.
  This tuning is examined by validation data.

	Validation accuracy
γ=1	0.625
γ=2	0.375
γ=3	0.5781

- AdaBoost (Using linear SVM)

	Train accuracy	Test accuracy
Before boosting	0.58125	0.5875
Boosting (iteration=2)	0.5625	0.4625
Boosting (iteration=5)	0.55625	0.525
Boosting (iteration=10)	0.63125	0.5625

2. PCA (Compress to 10 dimension)

	Train accuracy	Test accuracy
Bayesian Classifier	0.85	0.8
KNN (k=3)	0.828	0.8125
RBF kernel (σ=3)	0.8875	0.8875
Polynomial kernel (γ=1)	0.81875	0.8875

- KNN k tuning

	Train accuracy	Test accuracy
k=1	1	0.8125
k=3	0.828	0.8125

- RBF kernel σ tuning

RBF kernel SVM is optimized by gradient descent.

	Train accuracy	Test accuracy
σ=2	0.7656	0.7375
σ=3	0.8875	0.8875

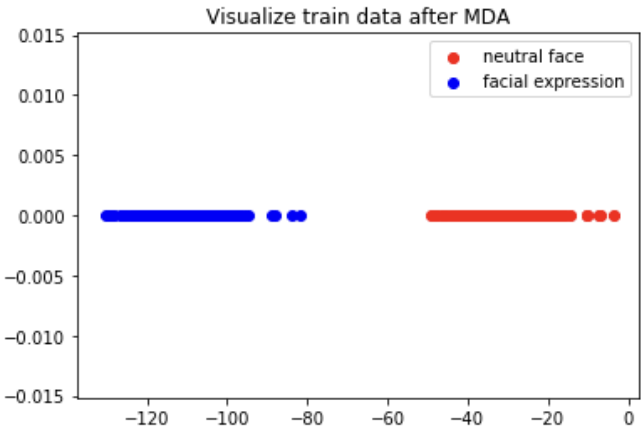
- AdaBoost (Using linear SVM)

	Train accuracy	Test accuracy
Before boosting	0.81875	0.8875
Boosting (iteration=10)	0.8625	0.825

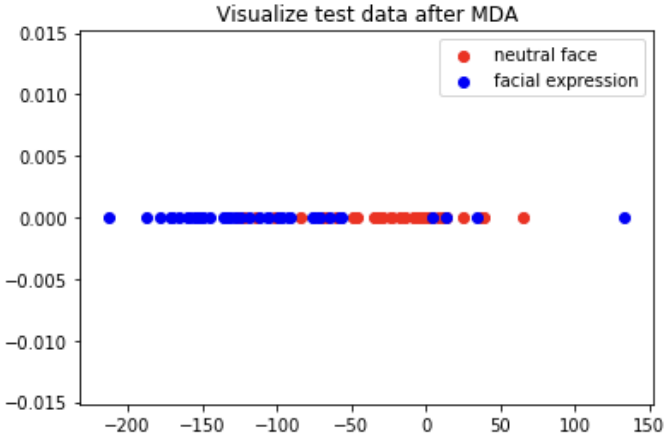
3. MDA

MDA is calculated with training data.

Training data distribution after MDA



Testing data distribution after MDA



	Train accuracy	Test accuracy
Bayesian Classifier	1	0.825
KNN (k=3)	1	0.7875
RBF kernel (σ=3)	1	0.7375
Polynomial kernel (γ=1)	0.5	0.575

- KNN k tuning

	Train accuracy	Test accuracy
k=1	1	0.7875
k=3	1	0.7875

Discussion:

Overall:

1. RBF kernel achieves the best overall accuracy with the original data.
2. As the k in KNN gets larger, the testing accuracy also increases. However, the train accuracy will drop because the train accuracy of k=1 will always be 1.
3. In the RBF σ tuning, σ = 3 has the highest accuracy
4. In the Polynomial kernel γ tuning, γ=1 has the highest accuracy. However, this result is more because of using gradient descent as an optimization tool. Higher γ kernel should be able to include the result of lower γ kernel. By using gradient descent, it requires fine tuning to get as close to the global optimal as possible.
5. From the results of AdaBoost, we can see the trends of increasing in the training accuracy. However, AdaBoost doesn't guarantee the improvement in test accuracy.

PCA:

After reducing to only 10 dimensions, all the calculations become very fast. And the accuracy remains over 80% for all the methods. This shows that most of the dimensions in the original data are useless in this classification task.

MDA:

Because this task is a binary classification task, the MDA is a normal LDA task.
From the graph train data after MDA, we can see that the data is separate from each other completely. Therefore, most of the classifiers can get 100% accuracy on the training accuracy.

However, the test accuracy drops compared with using PCA and original data. This is because most of the information is gone by projecting the data onto a single dimension. Therefore, the test data cannot be separated completely as the graphs show.

Polynomial kernel is the only classifier that didn't reach 100% training accuracy. This might be because of the non-optimal optimization with gradient descent. From the experiment, optimization with gradient descent of polynomial kernel is a lot harder compared to RBF kernel.