# Face Classification

Rahul Shah

April 2023

## 1 Introduction

These two Classifiers were performed for Subject Label Classification and Expression Classification:

- Bayes Classifier
- K-NN Classifier

To reduce the number of dimensions being fit to, these dimensional reduction techniques were used:

- PCA
- MDA

## 2 Data Sets

These three data sets were provided:

- Face - 200 subjects with 3 images each
  - First image is neutral
  - Second image has a facial expression
  - Third image has a different lighting condition
- Pose - 68 subjects with 13 images each under different poses
- Illumination - 68 subjects with 21 images each under different lighting conditions

## 3 Subject Label Classification

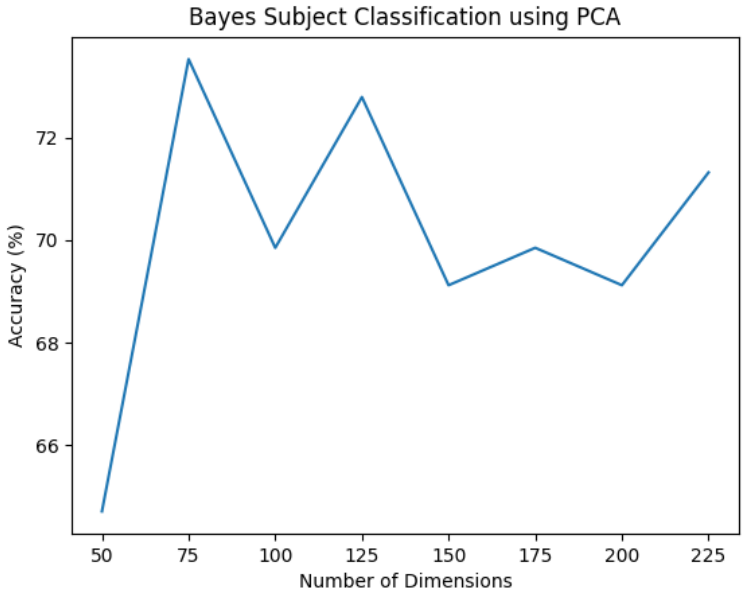
The POSE dataset was used. 4/5 of the dataset was used for training and the other 1/5 for testing. The purpose of this task is to identify the subject label from a test subject image.

### 3.1 Bayes

Only PCA was used to reduce the number of dimensions, MDA was not used.

First, PCA was performed to choose the number of dimensions to keep 95% of the data. This used between 180 and 190 dimensions, with an accuracy of above 60% for Bayes.

Second, PCA was performed again with a varying number of dimensions to see how the accuracy correlates with the number of dimensions kept for PCA.




This classifier is well above chance (50%). There is no clear correlation between number of dimensions used and accuracy. The accuracy is steady at approximately 70%. While the accuracy is not perfect with PCA, it is much more time-efficient than not using PCA. From this graph, 75-125 dimensions seem the most optimal as they produce resulsts with the highest accuracy.

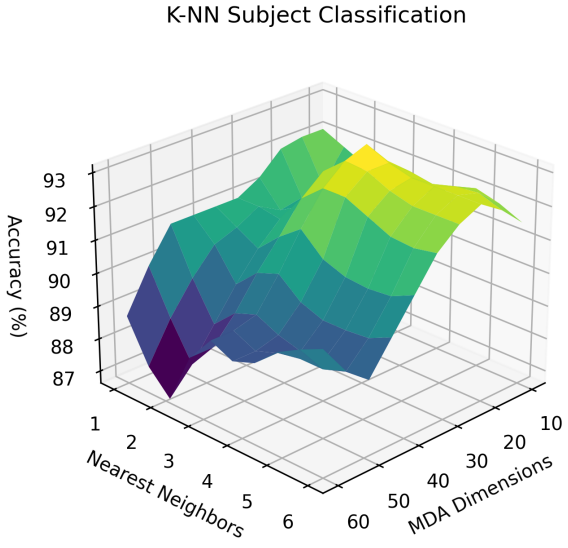
## 3.2 K-NN

PCA and MDA were used for K-NN.

### 3.2.1 MDA + PCA

PCA was run first to reduce dimensions to keep 95% of the data, then MDA was run with an arbitrary 20 number of dimensions. On an initial run With 3-NN, the accuracy was 93.7%.

To view correlation, the number of dimensions for MDA and K were varied to find the relation to accuracy. The number of dimensions for PCA were kept as optimal to avoid too much complexity.

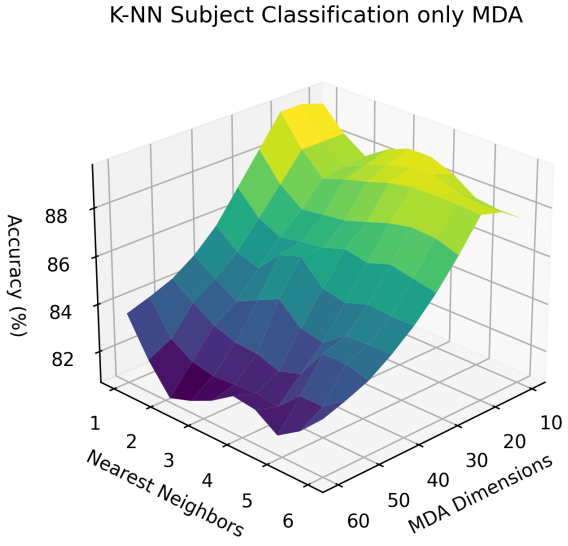



The accuracy is above 87% for all runs. It reaches a peak at ~93% with 3-NN and 10 dimensions. In

fact, as more MDA dimensions are used, the accuracy drops, reaching as low as ~87% with 2-NN. This is a result of over-fitting. More dimensions means more features of the training data are used. Some unimportant features are kept in the model and the testing data is classified with those as well. Keeping the complexity fairly low improves the accuracy here.

### 3.2.2 MDA

Here, only MDA was used to reduce the complexity of the data. The dimensions and K were still varied to show the new correlation for accuracy.




The accuracy drops to below 90% for all runs. The same correlation is there though, as the number of dimensions used in MDA increases, the accuracy decreases, dropping as low as ~82% with 3-NN.

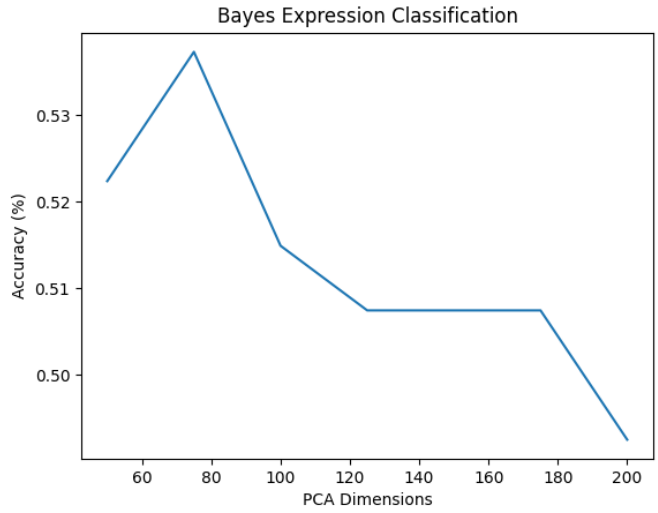
While PCA reduces the number of features of the data, it is still important to use it to remove the unimportant features and only classify the test data on the most important features.

# 4 Expression Classification

The FACE dataset was used. 2/3 was used for training and the other 1/3 for testing. The purpose of this task is to identify weather or not a test subject has a facial expression. There are two classes here (expresion or no expression).
PCA and MDA are used to include the optimal number of features for classification.
Because there are 2 classes, the number of dimensions for MDA is constant at 1.

## 4.1 Bayes

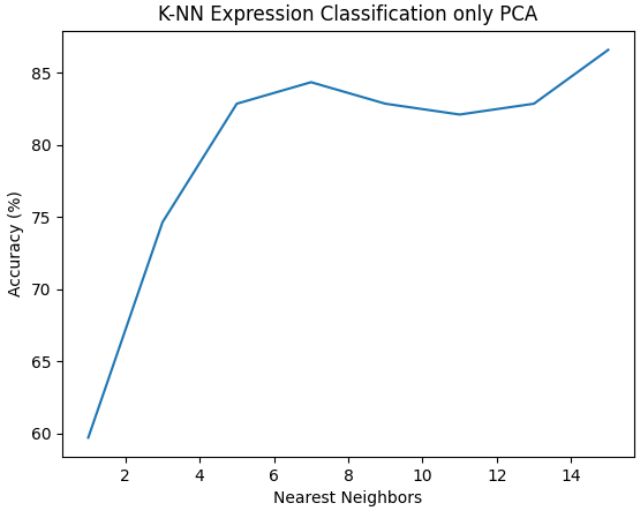
As usual, the number of dimensions for PCA is set to keep 95% of the features. This gives an accuracy of 62.6%, which is not great for a classifier.
The number of dimensions for PCA is varied to show correlation to accuracy.




This accuracy sits around chance level, which is not great for a Bayes classifier. Using PCA and MDA helped to reduce the complexity, but there may have been errors in the calculations for the classifier.

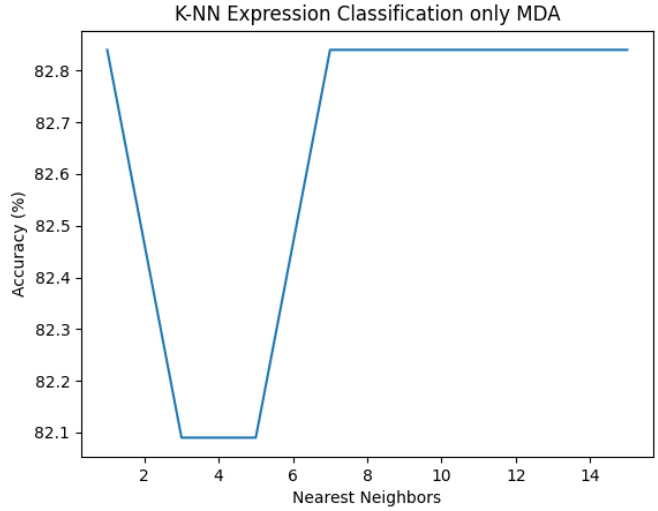
## 4.2 K-NN

### 4.2.1 PCA



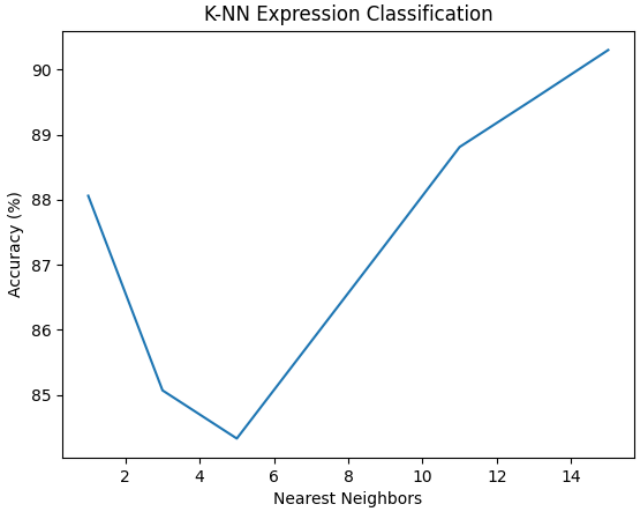

When only PCA was used, the accuracy increased and leveled off at ~85%. Increasing the number of NN was only beneficial until K= 6, then it started adding unnecessary complexity. This classifier kept the number of PCA dimensions optimal to preserve 95% of the data.

### 4.2.2 MDA




Only using MDA kept the accuracy steady at just under 83%, except for outliers from $K = 3 - 5$. There was overall little variance in the accuracy. Because there were only 2 classes, the number of dimensions had to stay at $d = 1$. So, only increasing K is not helpful in improving this accuracy

### 4.2.3 PCA + MDA




The accuracy started to decrease and then increase again as K increased. When $K = 15$. the accuracy reached 95% accuracy, however not much higher than 88% with $K = 1$.

## 5 Closing Thoughts + Errors

Increasing the complexity of the classifier may improve the accuracy, but not by much and would result in much higher completion times and possible over-fitting. It is important to be able to generalize the training data so that a variety of test data can be properly classified, not only those that fit certain criteria and forms.
K-NN performed much better than Bayes, most likely do it being less complex with less room for user error when creating the classifier. This was a major issue early on with the Bayes accuracy being almost 0% due to the predicted labels all giving out zero. After fixing this, there still could be hard to catch errors in the classifier.