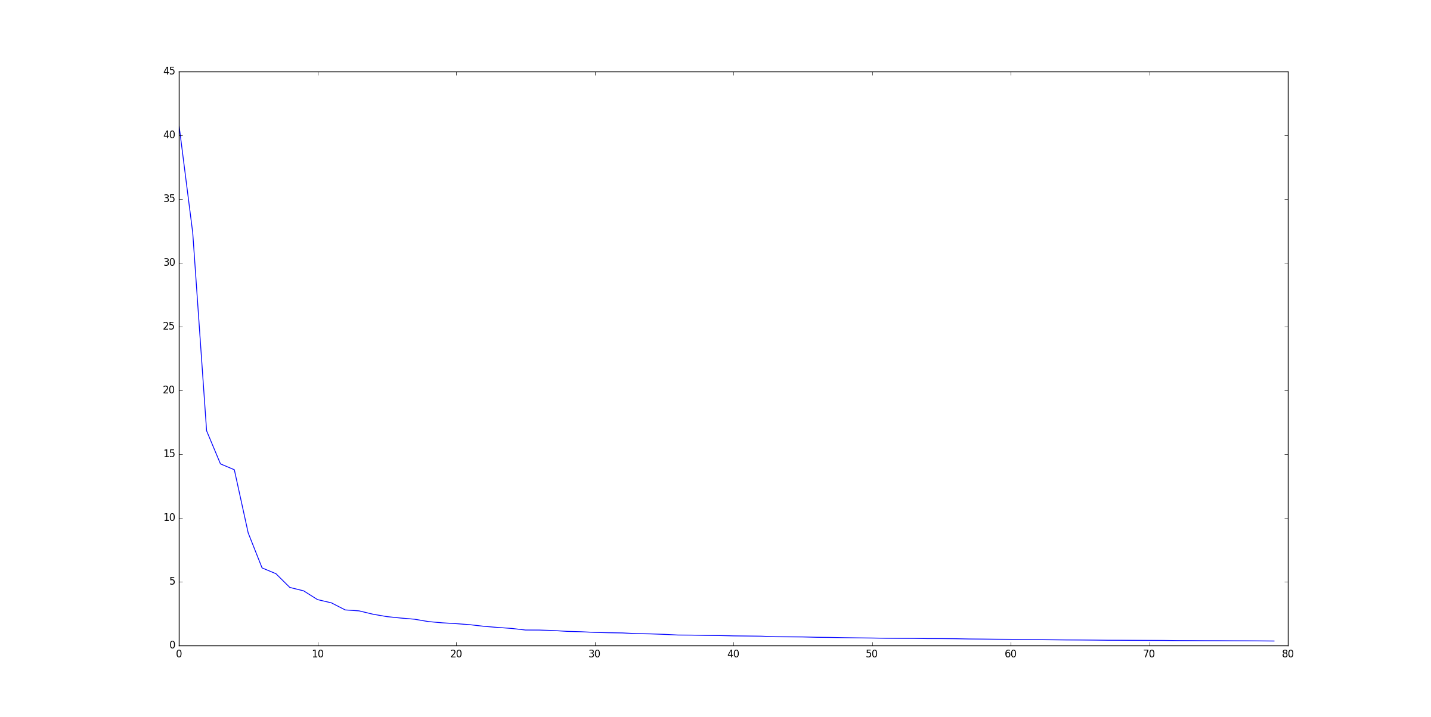
Report

On PCA LDA KNN and SVM

Task1:

Plot the eigenvalues here for PCA:



KNN using PCA for 5 Fold Cross Validation:

PCA = 80

|  |  |
| --- | --- |
| KNN using PCA for 1 Fold | 96.25 |
| KNN using PCA for 2 Fold | 95.0 |
| KNN using PCA for 3 Fold | 98.75 |
| KNN using PCA for 4 Fold | 97.5 |
| KNN using PCA for 5 Fold | 95.0 |
| Average | 96.5 |

Task 2:

Resize images from 112 x 92 to 56 x 46 and repeat Task 1, compare the new results to the results using un-resized images.

KNN Using PCA for 5-fold Cross Validation on Resized Images:

PCA = 80

|  |  |
| --- | --- |
| KNN using PCA for 1 Fold | 98.75 |
| KNN using PCA for 2 Fold | 98.75 |
| KNN using PCA for 3 Fold | 91.25 |
| KNN using PCA for 4 Fold | 100.0 |
| KNN using PCA for 5 Fold | 95.0 |
| Average | 96.75 |

**Task 3:**

LDA was applied on the dataset and top 39 Eigen values and corresponding Eigen vectors were selected. The dataset consists of 10304 features which are downscaled. The accuracy calculation for each fold with KNN is as follows:

|  |  |
| --- | --- |
| KNN using LDA for 1 Fold | 96.25 |
| KNN using LDA for 2 Fold | 100.0 |
| KNN using LDA for 3 Fold | 100.00 |
| KNN using LDA for 4 Fold | 95.00 |
| KNN using LDA for 5 Fold | 96.25 |
| Average | 97.5 |

Task 4:

PCA was applied on the dataset first and top 320 Eigen values and corresponding Eigen vectors were selected. Later LDA was applied and dimensions were further reduced to 39. The dataset consists of 10304 features which are downscaled. The accuracy calculation for each fold with KNN is as follows:

|  |  |
| --- | --- |
| KNN using LDA for 1 Fold | 96.25 |
| KNN using LDA for 2 Fold | 100.0 |
| KNN using LDA for 3 Fold | 92.50 |
| KNN using LDA for 4 Fold | 98.75 |
| KNN using LDA for 5 Fold | 96.25 |
| Average | 96.75 |

Tasks5:

Implement kernel SVM to do classification with 5-fold cross validation (any kernel is ok)

SVM Using PCA for 5 Fold Cross Validation:

PCA = 80

|  |  |
| --- | --- |
| SVM using PCA for 1 Fold | 100.0 |
| SVM using PCA for 2 Fold | 96.25 |
| SVM using PCA for 3 Fold | 97.5 |
| SVM using PCA for 4 Fold | 100.0 |
| SVM using PCA for 5 Fold | 95.0 |
| Average | 97.75 |

SVM Using PCA for 5 fold Cross Validation on Resized Images:

PCA = 80

|  |  |
| --- | --- |
| SVM using PCA for 1 Fold | 100.0 |
| SVM using PCA for 2 Fold | 96.25 |
| SVM using PCA for 3 Fold | 100.0 |
| SVM using PCA for 4 Fold | 98.75 |
| SVM using PCA for 5 Fold | 95.0 |
| Average | 98 |

**Task 6:**

PCA was applied on the dataset and top 80 Eigen values and corresponding Eigen vectors were selected. The dataset consists of 10304 features which are downscaled. The accuracy calculation for each fold with SVM is as follows:

|  |  |
| --- | --- |
| SVM using PCA for 1 Fold | 98.75 |
| SVM using PCA for 2 Fold | 100.00 |
| SVM using PCA for 3 Fold | 97.50 |
| SVM using PCA for 4 Fold | 98.75 |
| SVM using PCA for 5 Fold | 96.25 |
| Average | 98.45 |

Thus KNN is good choice when there are lot of points in a low dimensional space, while SVM is better when there are few points in higher dimensional space.