

ENEE633 PROJECT REPORT

Submitted by

Nithish Kumar Sanjeev Kumar
116316958

BAYES RULE FOR CLASSIFICATION:

Since priors are equal, maximum likelihood is proportional to the posterior probability. Higher the maximum likelihood higher the probability of being in a particular class.

Bayes rule for classification was applied with different preprocessing to compare its performance. The analysis on its accuracy and time consumption is presented below.

Method	Accuracy	Time Taken (s)
Bayes Classifier	72.5%	20.67
Bayes Classifier with PCA	75.56%	33.80
Bayes Classifier with LDA	80.46%	30.47

The above mentioned PCA and LDA are from the best possible configuration.

When Bayes classifier was utilized on the original mnist data, it gave a decent accuracy of 72%. But I needed more computation than when dimensionality was reduced and so the time needed was reduced. PCA and LDA are able to reduce the dimensionality still being reasonably accurate.

Analysis on the performance of PCA and LDA with respect to various parameters is shown in the next section.

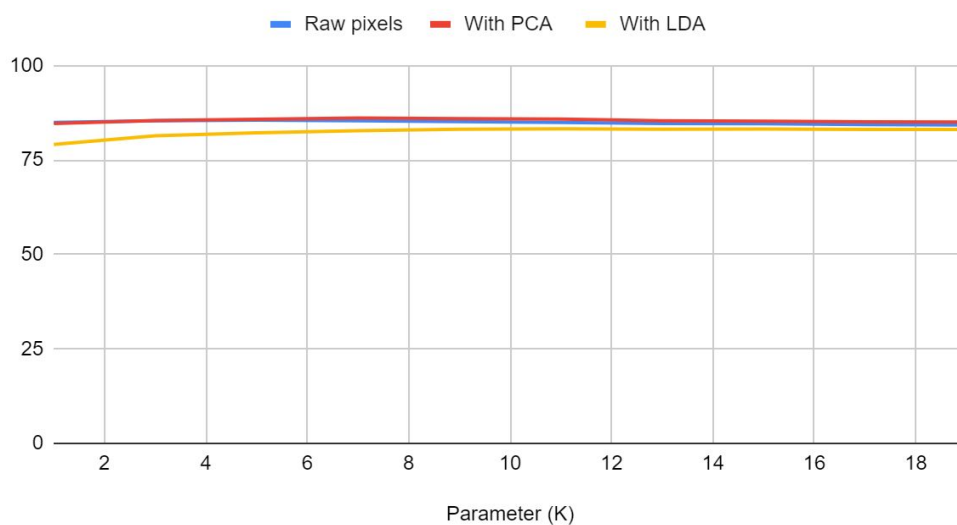
NEAREST NEIGHBOUR CLASSIFIER:

The core of this classifier depends on choosing the number of neighbours that is needed to make a decision on classification. This was analyzed and the results are presented.

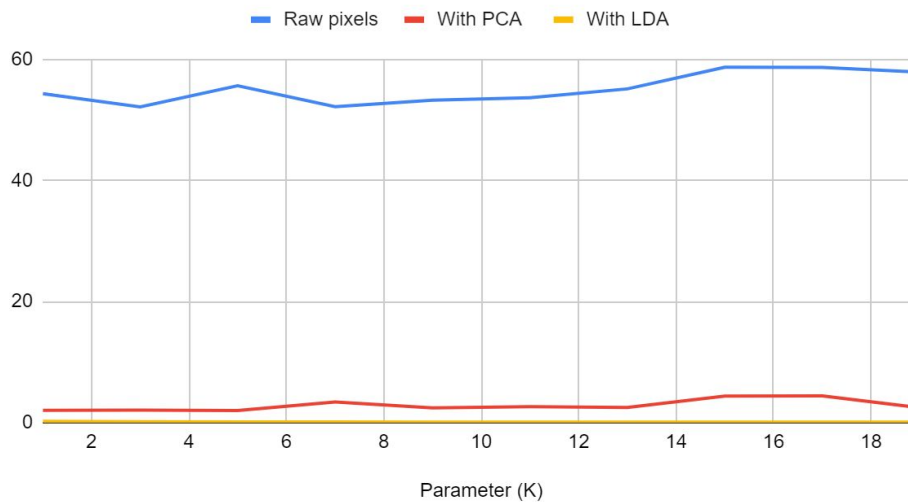
Results:

Parameter (K)	Raw pixels		With PCA (sigma =0.95)		With LDA(c = 9)	
	Accuracy	Time (s)	Accuracy	Time (s)	Accuracy	Time (s)
1	84.97	54.32	84.67	2.022	79.11	0.210615
3	85.41	52.15	85.43	2.07948	81.4	0.141145
5	85.54	55.63	85.78	2.01484	82.21	0.134274
7	85.40	52.19	86.11	3.40388	82.79	0.134714
9	85.19	53.26	85.97	2.45296	83.14	0.104064
11	84.95	53.67	85.82	2.62575	83.25	0.0735228
13	84.68	55.12	85.39	2.51105	83.16	0.0733624
15	84.62	58.69	85.28	4.39536	83.22	0.0718098
17	84.41	58.65	85.09	4.41091	83.1	0.0747006
19	84.27	57.89	85.05	2.47696	83.09	0.078666

ACCURACY



Time taken (s)



Analysis:

For the raw data, nearest neighbour classifier with $k=5$ has the highest accuracy of 85.54%. This is a good improvement compared to the bayes classifier. Varying the number of neighbours doesn't seem to have a great impact on this dataset. Using PCA with a covariance of 0.95 maintained the same average accuracy. But with LDA the average accuracy reduced by a bit but for a very great reduction in computation time.

PRINCIPAL COMPONENT ANALYSIS:

Principal Component Analysis (**PCA**) is **used to** explain the variance-covariance structure of a set of variables through linear combinations. It is commonly used dimensionality-reduction technique.

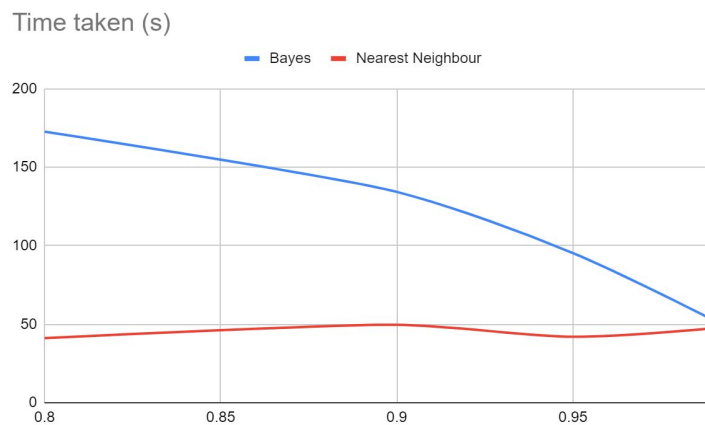
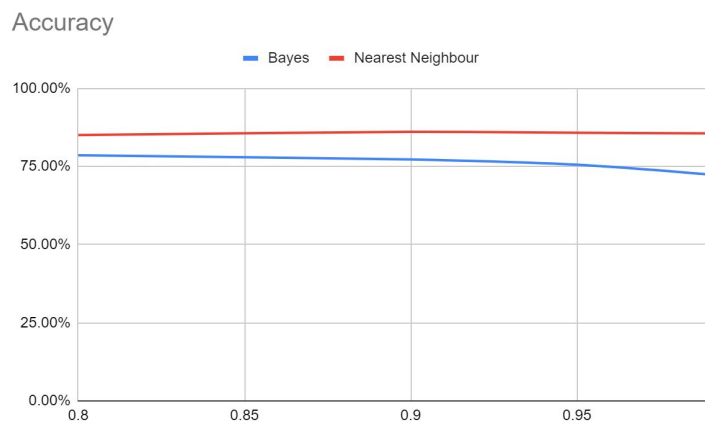
Variance is the parameter that helped in maintaining the accuracy in dataset. If the variance is allowed to be changed greatly then the data is reduced to a dimension that doesn't provide enough useful information.

Variation of variance and the corresponding accuracy in bayes classifier and nearest neighbour classifier is listed below.

Variance retained	Number of components	Bayes		Nearest Neighbour	
		Accuracy	Time (s)	Accuracy	Time
.99	527	72.4%	52.4	85.60%	47.43
.95	256	75.56%	95.3	85.78%	42.00
.9	137	77.25%	134.3	86.08%	49.64
.8	50	78.56%	172.7	85.06%	41.13

*Time includes dimensionality reduction time and computation time

As the variance retained is high then the data is more accurate and so is the classification. While the time taken to reduce it to a lower dimension is significantly higher as the variance change is higher.



Linear Discriminant Analysis

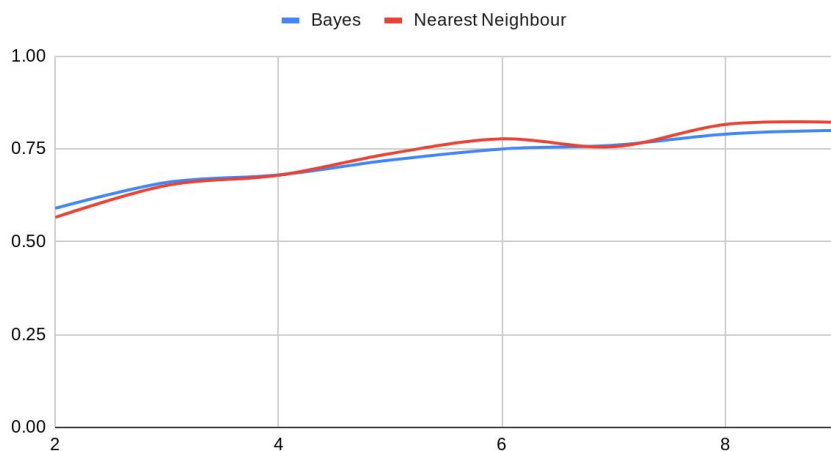
LDA being another common dimensionality-reduction technique projects the higher dimensional data on lesser number axes. In a dataset containing n-classes, the dimensionality can be varied from 1 to n-1.

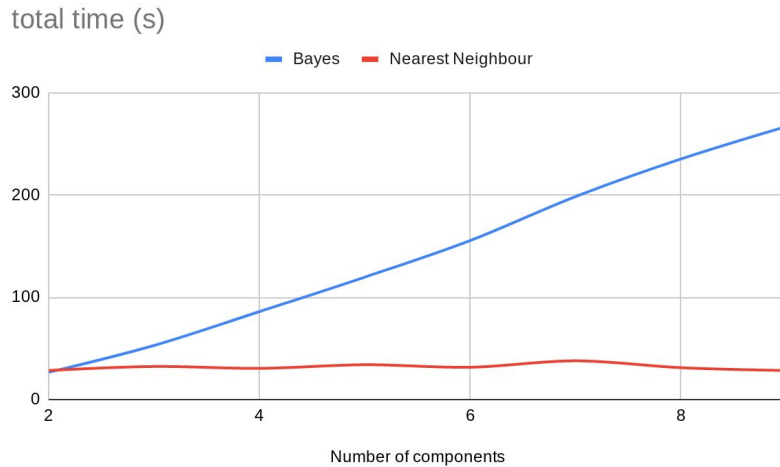
The trends in the behavior is listed below.

Number of components	Bayes		Nearest Neighbour	
	Accuracy	Time (s)	Accuracy	Time (s)
2	59%	26.62	56.57%	28.37
3	66%	52.88	65.19%	32.31
4	68%	85.98	67.92%	30.47
5	72%	119.715	73.75%	34.02
6	75%	155.53	77.73%	31.55
7	76%	198.50	75.6%	37.85
8	79%	235.28	81.61%	31.11
9	80%	267.03	82.21%	28.37

*Time includes reduction time and computation time

Accuracy





LDA reduces the dimensionality better and makes the computation time lesser. But it takes more time than PCA to reduce the dimension of the data. But this means the computation time itself is less, but the total time might seem like its greater than PCA, but it's not.

It seems to have done a good job wrt maintaining the accuracy in this dataset.

CONCLUSION:

Based on the experiments performed Nearest neighbour classifier seems to be more accurate in classifying this fashion mnist dataset by a small margin. PCA and LDA reduces the dimensionality of the dataset and reduces the computation time greatly. LDA reduces the size of the dataset greatly while maintaining the similar accuracy. On the other hand, PCA's data is more dense than the one from LDA and could be used to classify much difficult test data. Use of PCA and LDA should be a tradeoff between the accuracy expected in the output and computation time available.