Pattern Recognition

CS6690

IIT Madras

Assignment 2

By:
Group 25:
Arun Baby (CS15S016)
Vishal Subbiah (MM12B035)

Bayesian Classifiers

In this assignment we build a baysian classifier to divide the data in to different classes. We used the bayes theorem for classification. The bayes probability is calculated as follows :-

$$P(w_i/\bar{x}) = \frac{p(\bar{x}/w_i) * P(w_i)}{p(\bar{x})} \tag{1}$$

Experiments

For the dataset, 70% of the data are considered for training and 30% of the data are considered for testing. There are 5 cases each for different configuration of covariance matrices. The different datasets include artificial and real data.

We assume all the data follows a gaussian distribution. We also assumed that $P(w_i)$ (prior probability) is equal among all classes. The two variables we need to estimate are mean μ and covariance C.

$$\hat{\mu} = \sum_{i=1}^{N} \frac{\bar{x}_i}{N} \tag{2}$$

$$C = \hat{\sigma^2} = \sum_{i=1}^{N} \frac{(\bar{x}_i - \hat{\mu})(\bar{x}_i - \hat{\mu})^T}{N-1} \tag{3}$$

$$p(\bar{x}/w_i) = \frac{1}{(2\pi)^{d/2}|C_i|^{1/2}} exp(-\frac{1}{2}(\bar{x} - \bar{\mu_i})^T C^{-1}(\bar{x} - \bar{\mu_i})) \tag{4}$$

The different cases we tried are :

Case 1: Bayes with Covariance same for all classes :- $C_1 = C_2 = ... = C_K$

Case 2: Bayes with Covariance different for all classes :- $C_1 \neq C_2 \neq C_K$

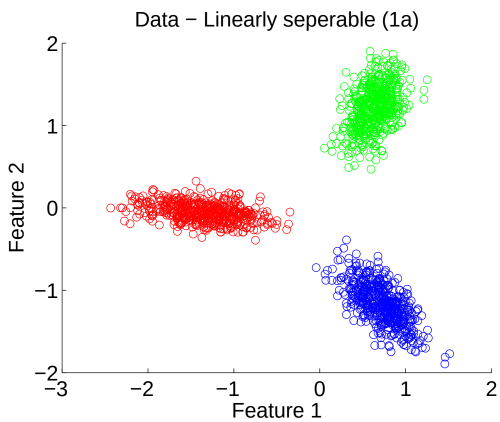
Case 3: Naive Bayes :- $C_1 = C_2 = C_K = \sigma^2 * I$

Case 4: Naive Bayes with C same for all :- $C = C_1 * I = C_2 * I = C_K * I$

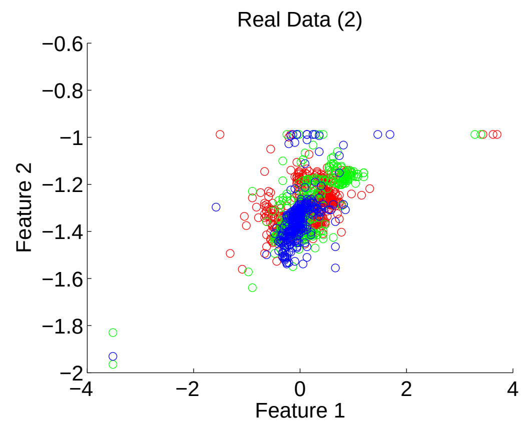
Case 5: Naive Bayes with C different for all :- $C_1 * I \neq C_2 * I \neq C_K * I$

Results

The following plots represent our results for the different data sets.



(a) 2-dimensional artificial data of 3 classes that are linearly separable

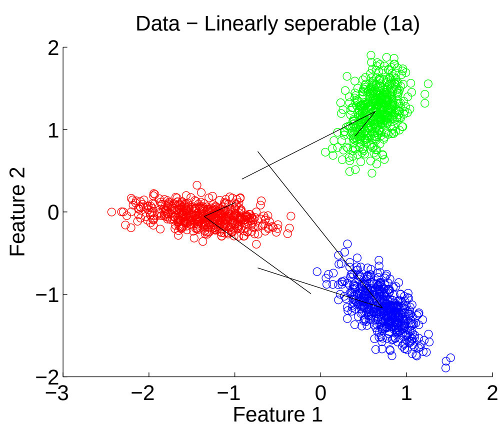


(b) Real world data of 3 classes

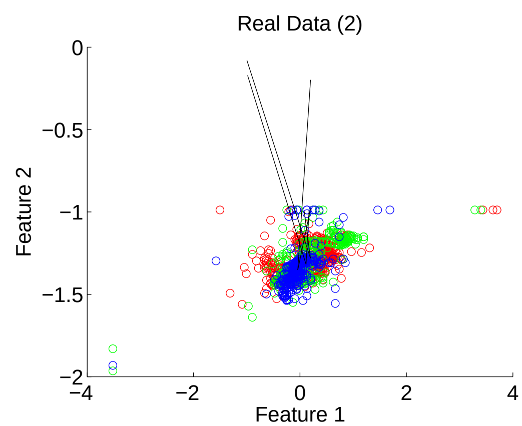
Number of classes: 3
Number if Features: 2
Number of Training samples: 350, 350, 350
Number of testing samples: 150, 150, 150

Number of classes: 3
Number if Features: 2
Number of Training samples: 1742, 1672, 1604
Number of testing samples: 746, 716, 687

Figure 1: Given Data sets

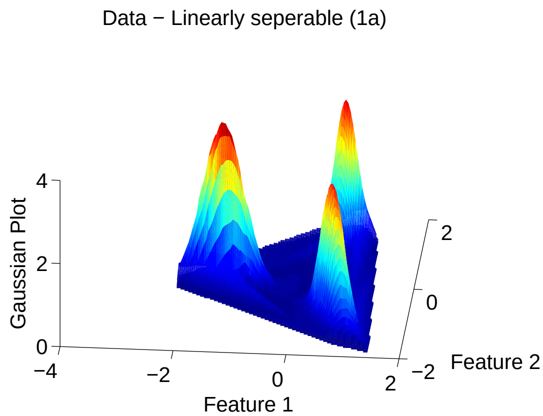


(a) eigen vectors of the 2-dimensional artificial data of 3 classes that are linearly separable

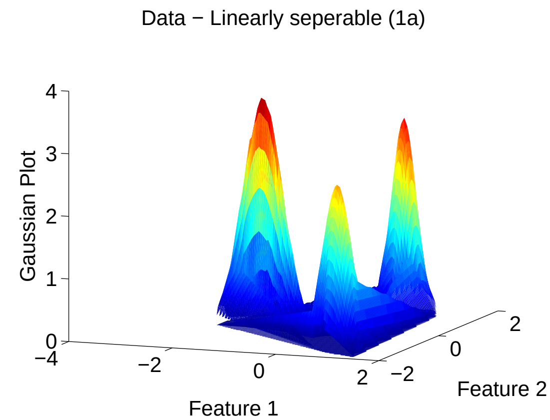


(b) Eigen vectors of Real world data of 3 classes

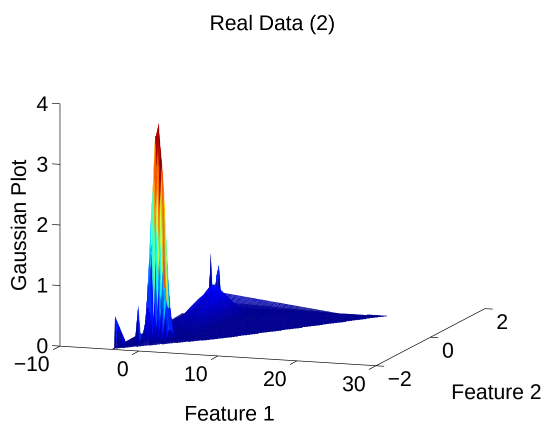
Figure 2: Eigen vectors of the given data



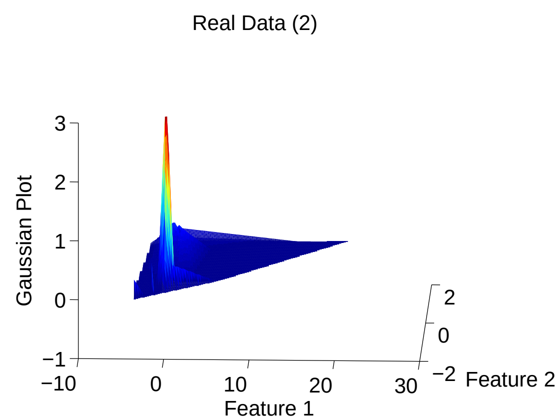
(a) Guassian plot of the 2-dimensional artificial data of 3 classes that are linearly separable for Case 2



(b) Guassian plot of the 2-dimensional artificial data of 3 classes that are linearly separable for Case 5



(c) Guassian plot of the Real world data of 3 classes for Case 2



(d) Guassian plot of the Real world data of 3 classes for Case 5

Figure 3: Gaussian plots

	Class One	Class Two	Class Three
Class One	742	2	2
Class Two	3	694	19
Class Three	3	86	598

Table 1: Confusion Matrix: Real world data for Case 3

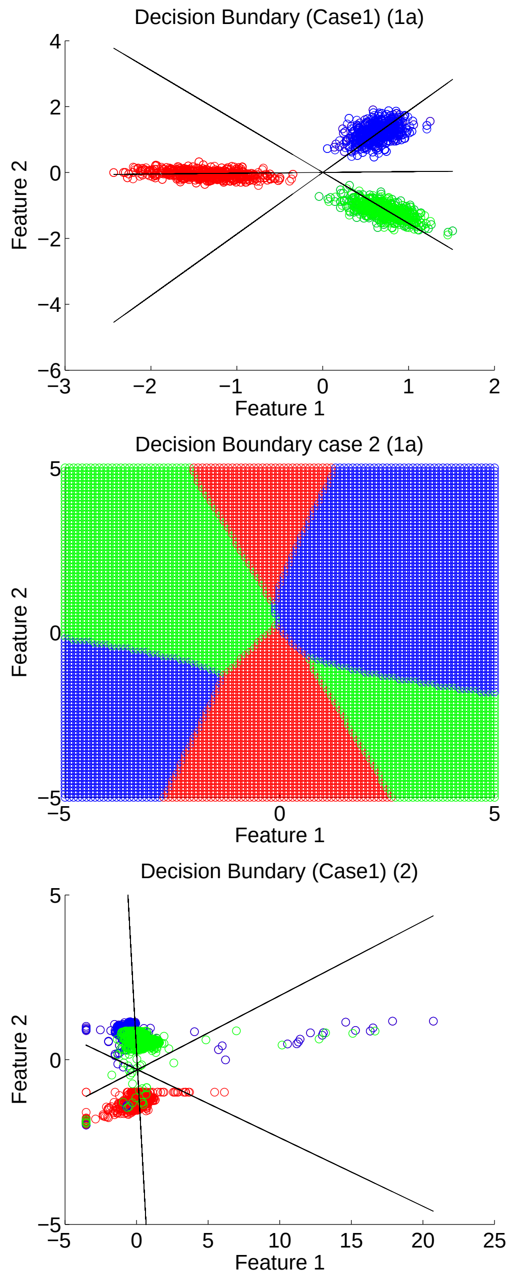


Figure 4: Decision Boundaries

	Class One	Class Two	Class Three
Class One	150	0	0
Class Two	0	150	0
Class Three	0	0	150

Table 2: Confusion Matrix: Artificial data which is linearly separable for Case 2

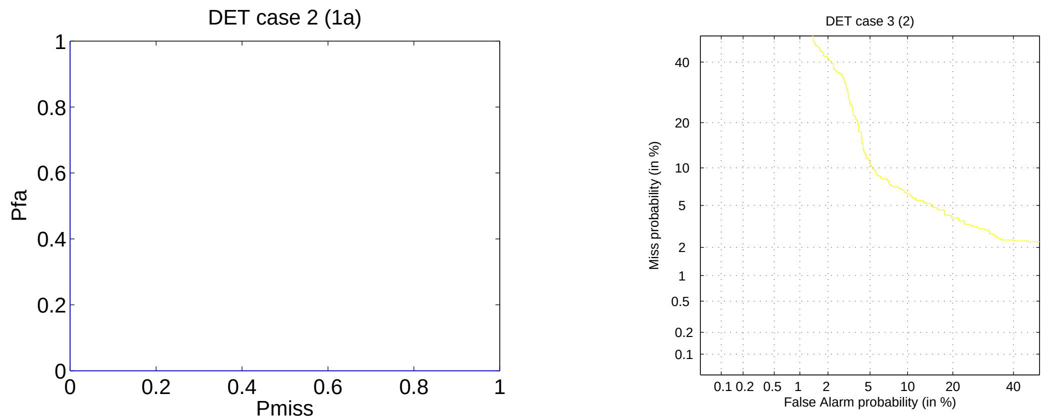


Figure 5: Detection Error Tradeoff

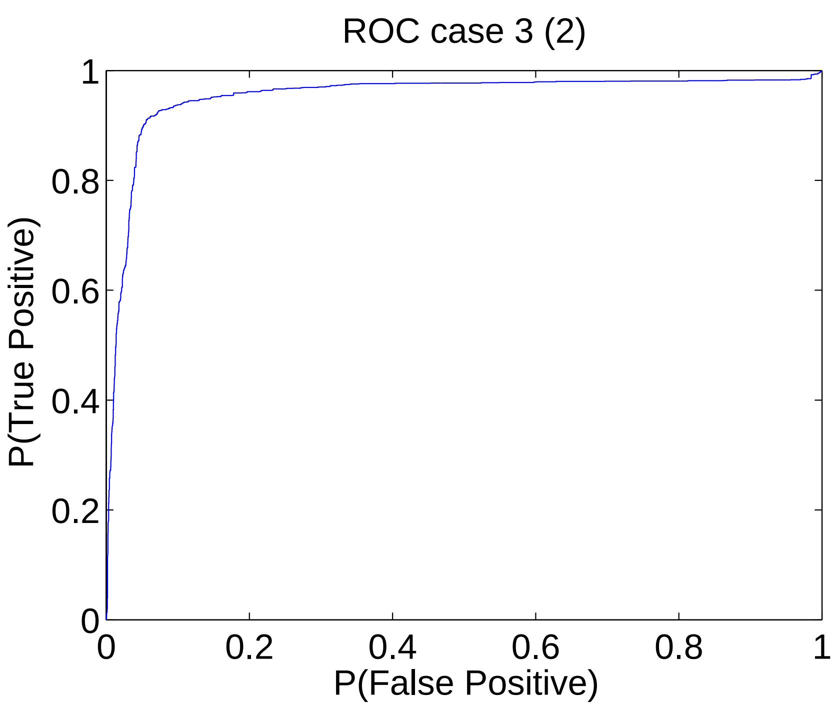


Figure 6: Receiver Operating Characteristic

Observations

Case	1	2	3	4	5
Accuracy	100	100	100	100	100

Table 3: 2-dimensional artificial data of 3 classes that are linearly separable

Case	1	2	3	4	5
Accuracy	100	100	100	100	100

Table 4: 2-dimensional artificial data of 3 classes that are nonlinearly separable

Case	1	2	3	4	5
Accuracy	99.11	98.89	99.11	99.11	99.89

Table 5: 2-dimensional artificial data of 3 classes that are overlapping

Case	1	2	3	4	5
Accuracy	94.6952	82.5035	94.6487	94.6487	91.7636

Table 6: Real world data of 3 classes

- By scaling the data the accuracy improves.
- The Bayes and Naïve Bayes classification results in high performance if mean of the distribution is well separated.
- Alignment of covariance matrices eigenvectors played major role in classification of features.
- Applying SVD on the covariance matrix, returned similar results for case 1, 3 and 4, while for case 2 and 5 it did not return as good results.