Suggested Lab 1 Report Outline

Cover Page

* Signed on the soft and hard copy

Introduction

* Brief explanation of what the purpose of the experiment

Theory

* Derives bayes theorem, explains its significance and the meaning of each variable (prior, scaling factor, likelyhood)
* Theory behind error, and choosing the decision boundary
* **How bayes decision rule is implemented in our two category case dichomomizer**

Results

* Graphs, plots

Discussion

* Answers to discussion questions

Conclusion

* Brief summary of purpose and what was collected and found

Results

Below are the following plots produced when runlab1.m is executed:

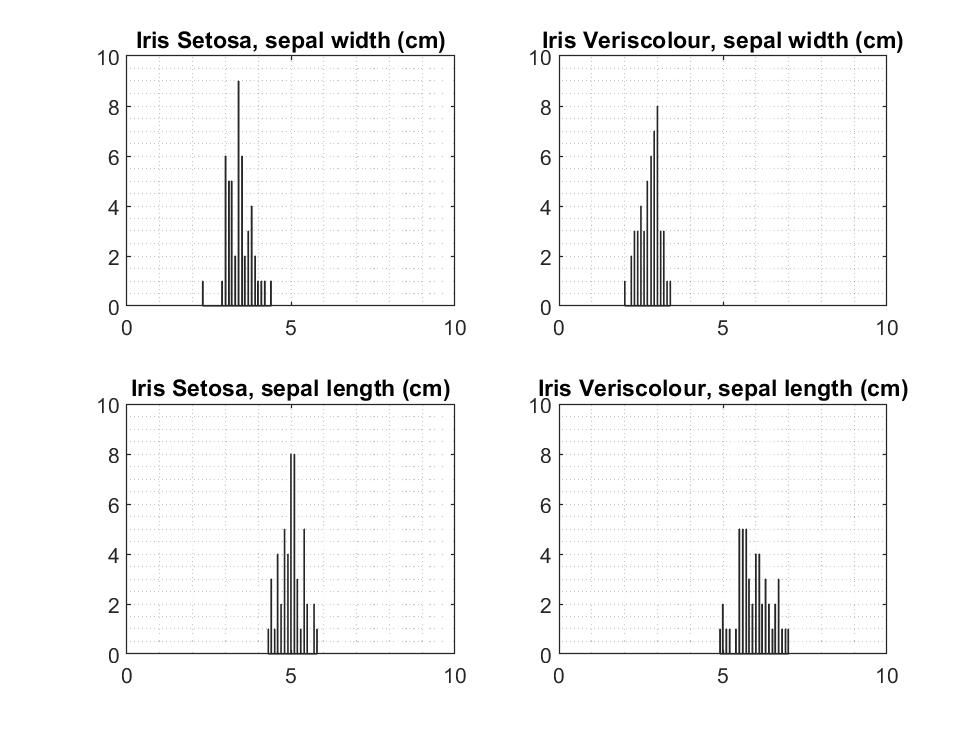


Figure #: Individual Histograms

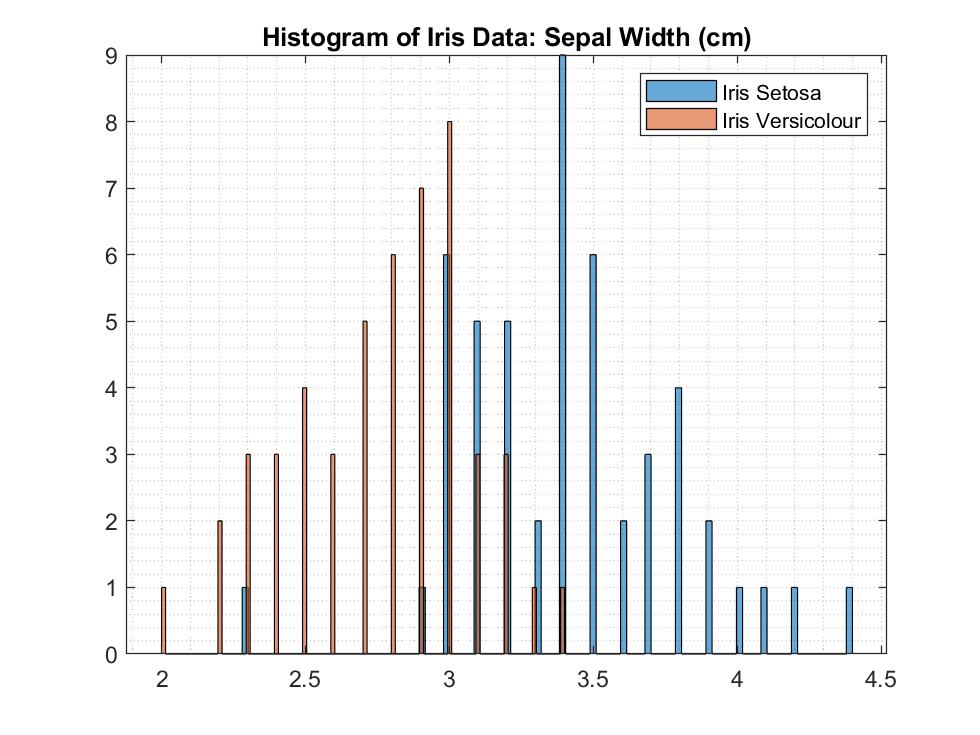


Figure #: Joint histogram showing sepal width data for Iris Setosa and Iris Versicolour classes

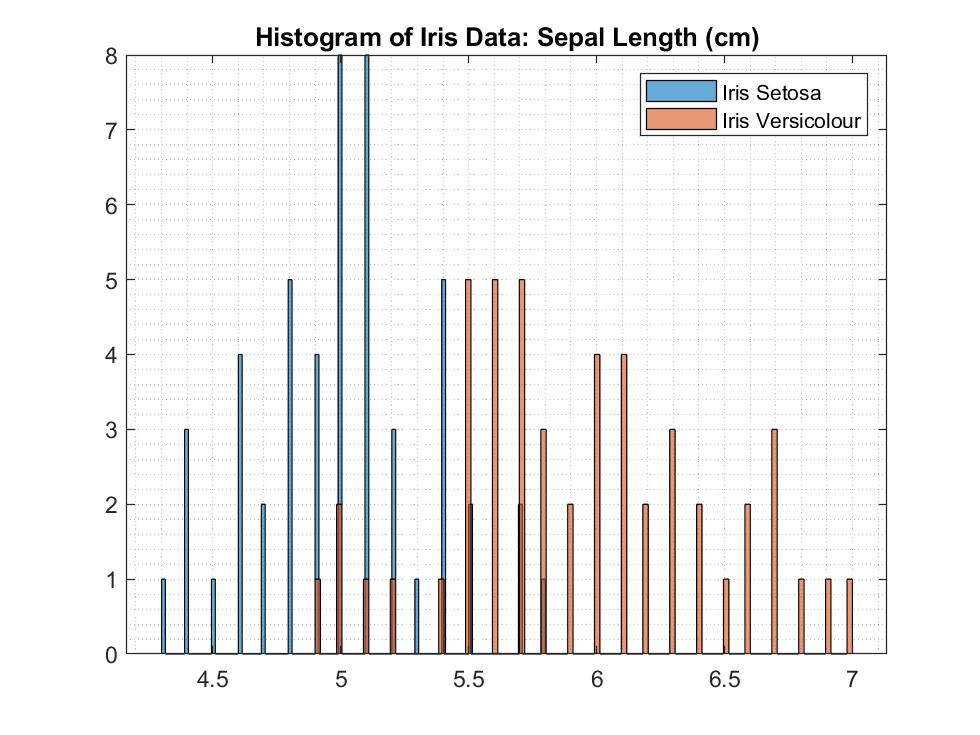
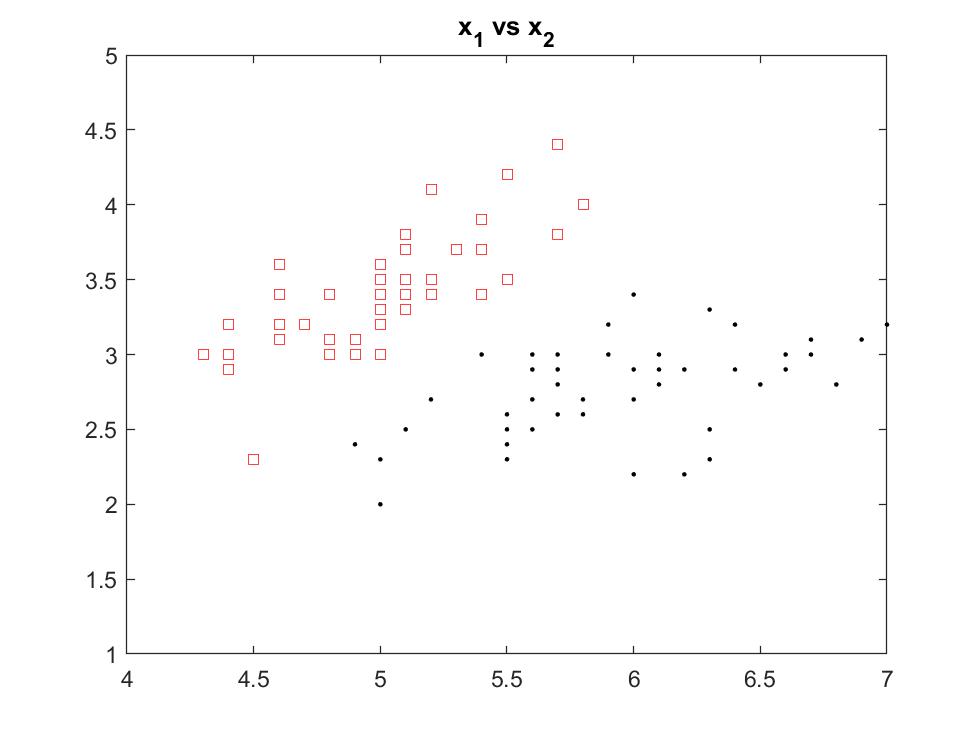


Figure #: Joint histogram showing sepal length data for Iris Setosa and Iris Versicolour classes



**Figure #**: Distribution of feature x1 and x2 for Iris Setosa and Iris Versicolour classes

Identifying the classes and their respective posteriors probabilities and discriminant function values with the input vector x = [3.3, 4.4, 5.0, 5.7, 6.3]

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Input value (x) | Posterior g1(x1) | Posterior g2(x1) | Discriminant value g(x1) | Class label Decision |
| **3.3** | 0.2753e-05 | 0.0730e-05 | 2.0232e-06 | Iris Setosa |
| **4.4** | 0.0766 | 0.0040 | 0.0726 | Iris Setosa |
| **5.0** | 0.3359 | 0.0649 | 0.2710 | Iris Setosa |
| **5.7** | 0.0484 | 0.3026 | -0.2543 | Iris Versicolour |
| **6.3** | 0.0004 | 0.2620 | -0.2616 | Iris Versicolour |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Input value (x) | Posterior g1(x2) | Posterior g2(x2) | Discriminant value g(x2) | Class label Decision |
| **3.3** | 0.3080 | 0.0776 | 0.2304 | Iris Setosa |
| **4.4** | 0.0117 | 0.0000 | 0.0117 | Iris Setosa |
| **5.0** | 0.5835e-04 | 0.0000 | 5.8350e-05 | Iris Setosa |
| **5.7** | 0.5253e-08 | 0.0000 | 5.2533e-09 | Iris Setosa |
| **6.3** | 0.1219e-12 | 0.0000 | 1.2193e-13 | Iris Setosa |

Decision boundary threshold (Th1) for sepal length feature x1 is approximately 5.38332 cm.

Decision boundary threshold (Th2) for sepal width feature x2 is approximately 3.06266 cm.

Discussion

The decision boundary threshold values were obtained through trial and error until posterior g1(x) is approximately equal to posterior g2(x). This is justified through Bayes Decision rule and a two-case category dichotomizer.

Bayes Decision Rule states:

Decide w1 if g(x) > 0; otherwise decide w2

Since the computation of the single discriminant function g(x) is as follows

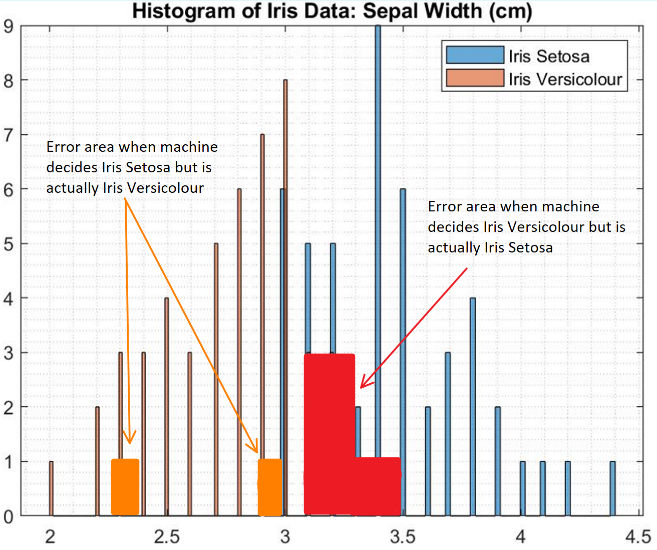
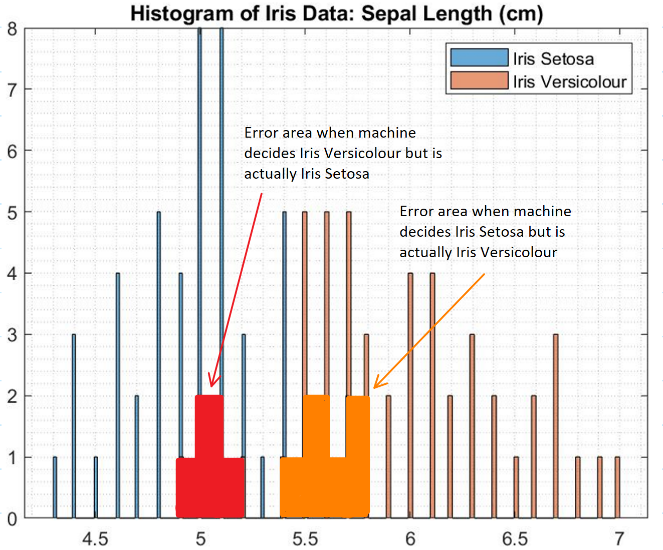
g(x) = g1(x) – g2(x)

the decision boundary will be a value at which any data greater than or less than will result to a class decision. Therefore theoretically, the threshold value should produce posteriors g1(x) and g2(x) that are equal which should then result in g(x) = 0.

The decision boundary is most often task and cost specific and can change depending on the penalty associated with misclassification. Looking at the joint histogram of sepal width in **Figure #**, if a higher penalty is associated with misclassifying w2 as w1 (red area), where w1 = Iris Setosa and w2 = Iris Versicolour, then the decision boundary can be adjusted (ie Th2 = 3.5 cm) such that there is little error with misclassification.

From **Figure #,** the feature sepal length has an error probability with approximately equal weight on either side of the threshold. On the other hand, the feature sepal width has an uneven error probability, where it is more likely to misclassify w2 as w1. Therefore, in deciding which feature to use for classification would depend on the task or cost at hand. In a more general sense, it would be preferred to use sepal length as a feature to classify due to equal distribution of error probability however, if let us say a situation is presented where there is a higher penalty incurred when w2 is misclassified as w1, then the sepal width is a more desirable feature for classification.

One of the goals with pattern recognition is to minimize error as much as possible—this can be done by gathering and processing data that can clearly discriminate between classes. Using two features for classification can give a clearer and more discriminating data compared to a single feature. **Figure #** displays better data when classifying w1 and w2 since there is little to no overlapping datapoints and shows a clear area of separation.



**Figure #:** Joint Histograms showing error area

Conclusion

In conclusion, it was found that the threshold Th1 for feature x1 is approximately 5.38332 cm, and the threshold Th2 for feature x2 is approximately 3.06266 cm. These decision boundaries were theoretically derived using trial and error from Bayes Decision rule and a two-case category dichotomizer. These threshold values are most often task or cost specific and thus allowed to be adjusted depending on the heaviness of the penalties incurred from misclassifying. Lastly, using a single feature to classify can be inadequate and most often hold larger error probabilities, thus it was found that using two features for classification can produce a data set that can be used to better discriminate between the classes.