Experimental Results:

<u>**Linear data**</u> <u>**Non-linear data**</u> <u>**Real-world data**</u>

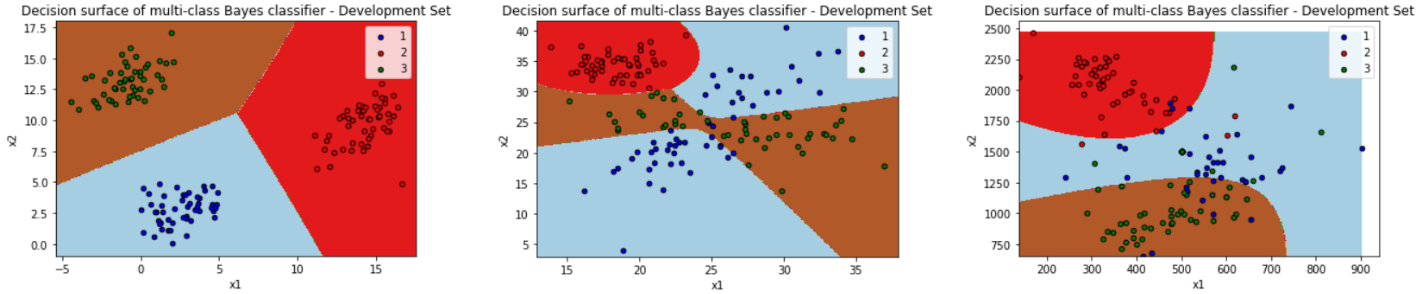

Figure 1: Decision boundary for a Bayes classifier with different covariance for the classes

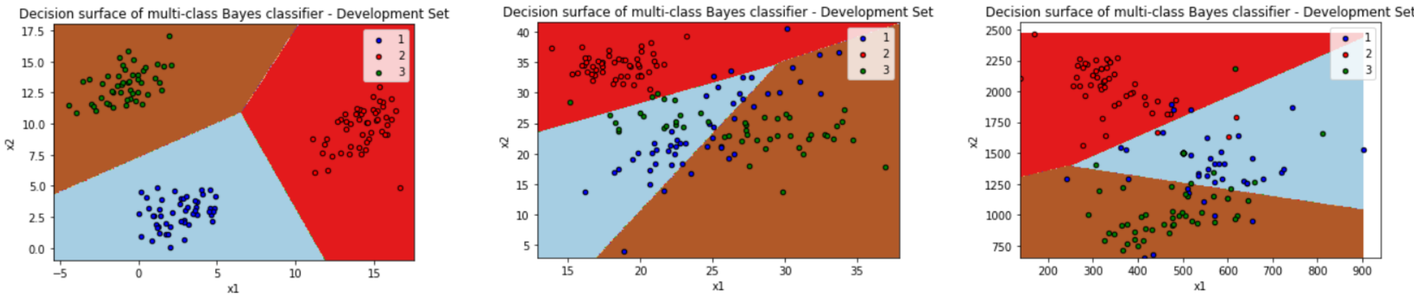

Figure 2: Decision boundary for a Bayes classifier with the same covariance for the classes

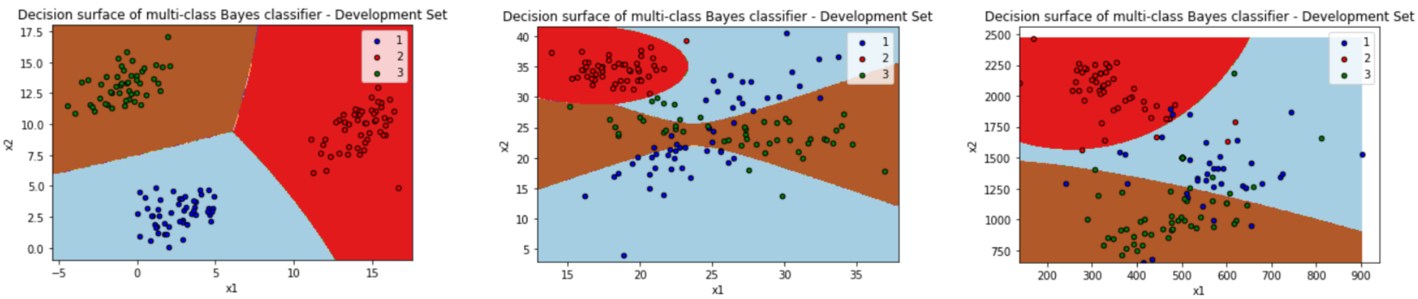

Figure 3: Decision boundary for a Naive Bayes classifier with different covariance for the classes

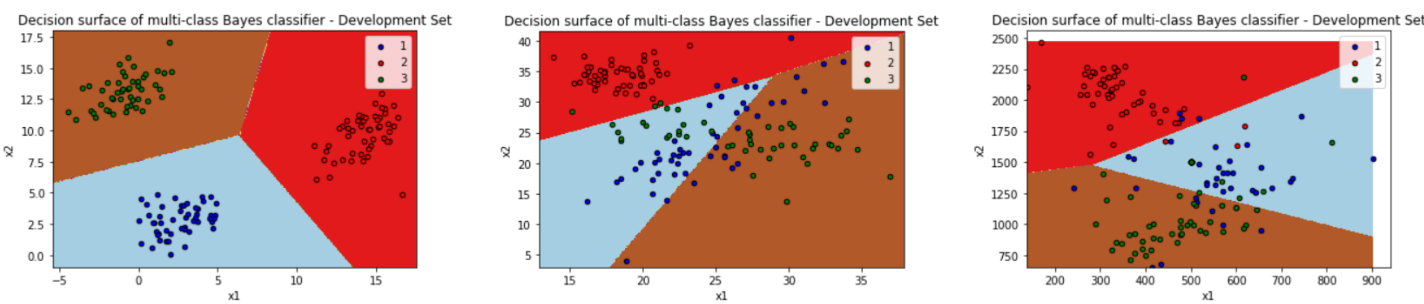

Figure 4: Decision boundary for a Naive Bayes classifier with the same covariance for the classes

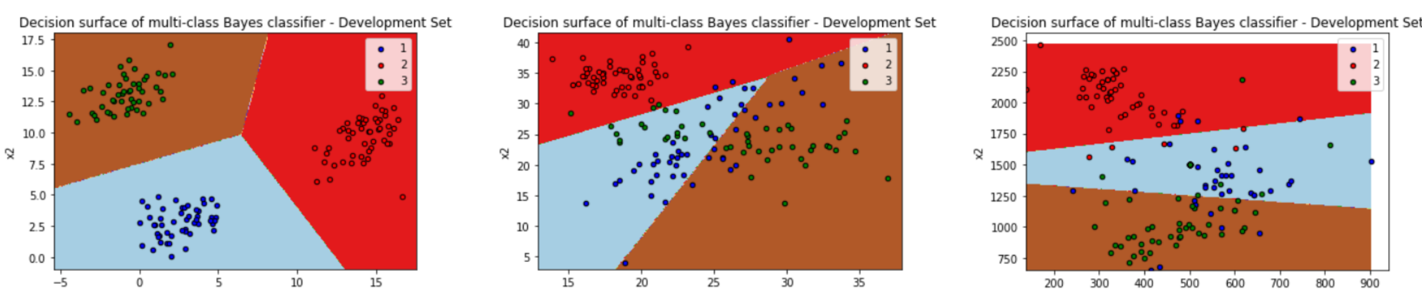

Figure 5: Decision boundary for a Naive Bayes classifier with $\Sigma = \sigma^2 I$

Linear data **Non-linear data** **Real-world data**

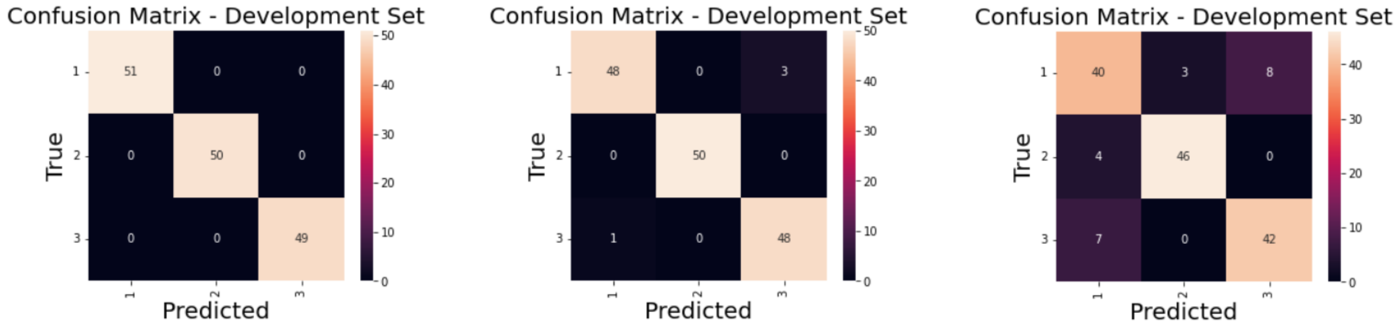

Figure 6: Confusion matrices for a Bayes classifier with different covariance for the classes

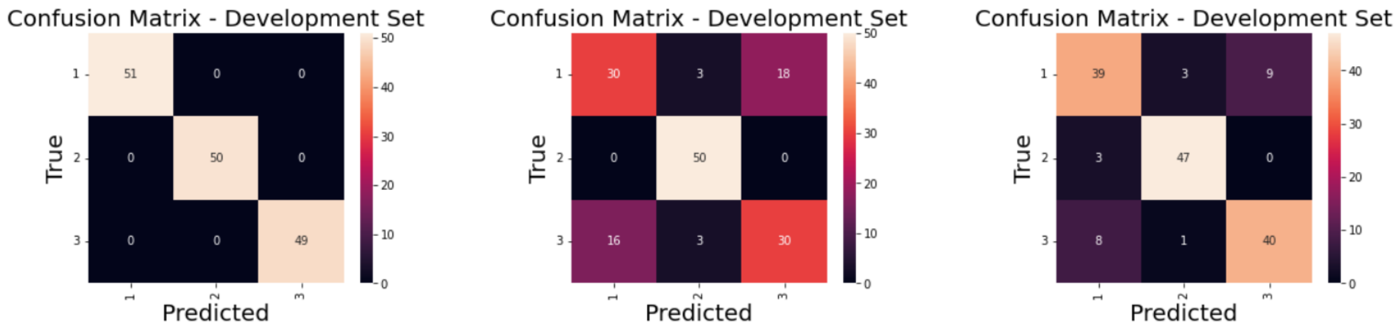

Figure 7: Confusion matrices for a Bayes classifier with the same covariance for the classes

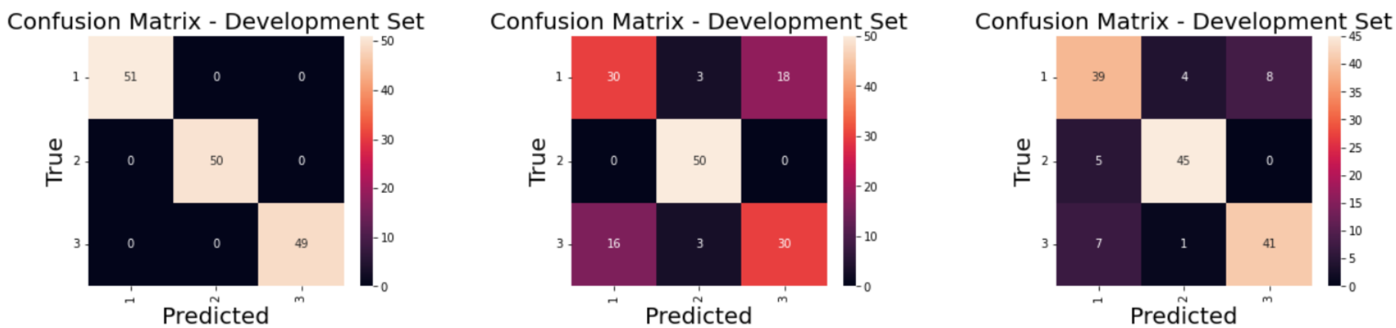

Figure 8: Confusion matrices for a Naive Bayes classifier with different covariance for the classes

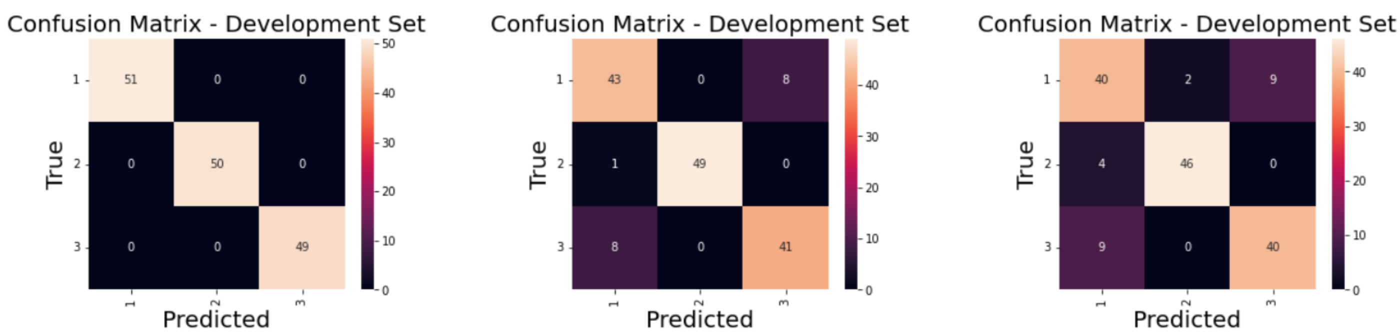

Figure 9: Confusion matrices for a Naive Bayes classifier with the same covariance for the classes

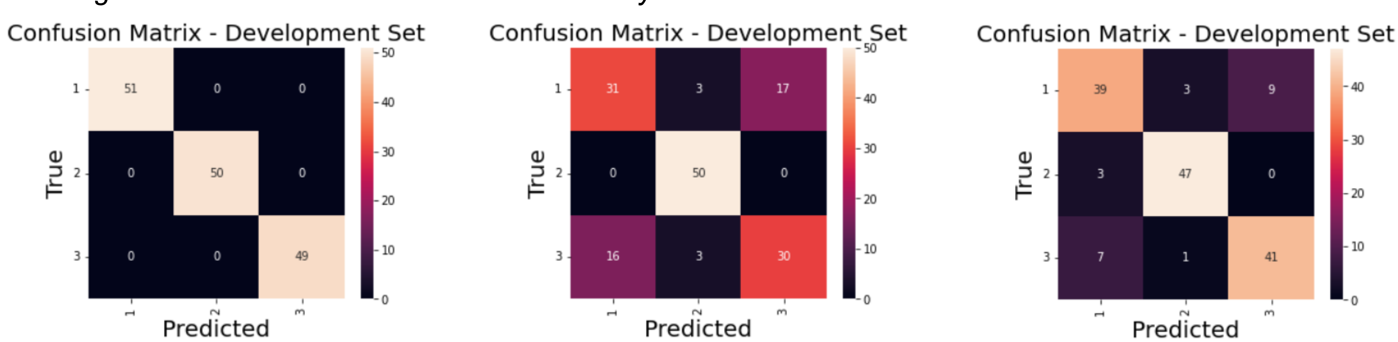

Figure 10: Confusion matrices for a Naive Bayes classifier with $\Sigma = \sigma^2 I$

Inference:

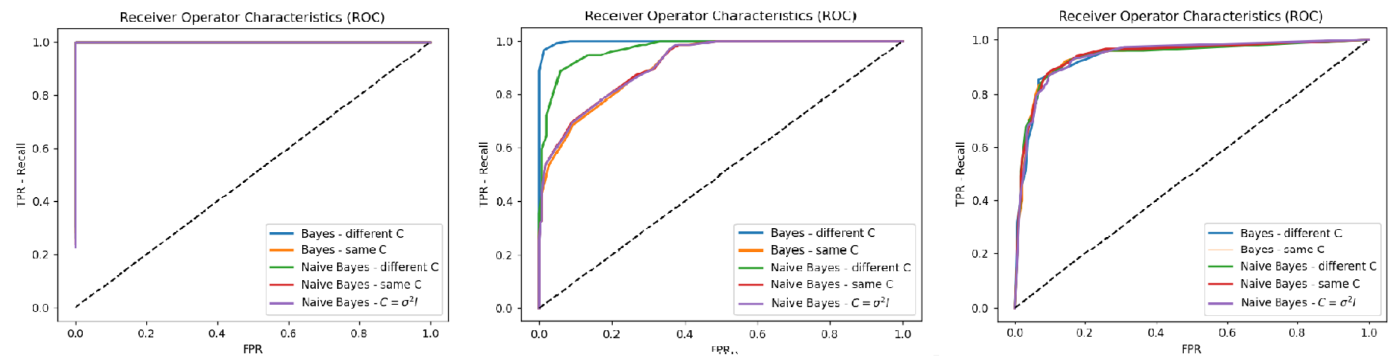

Figure 11: ROC curves of all the classifiers for linear data (left), non-linear data (middle), and real-world (right) data

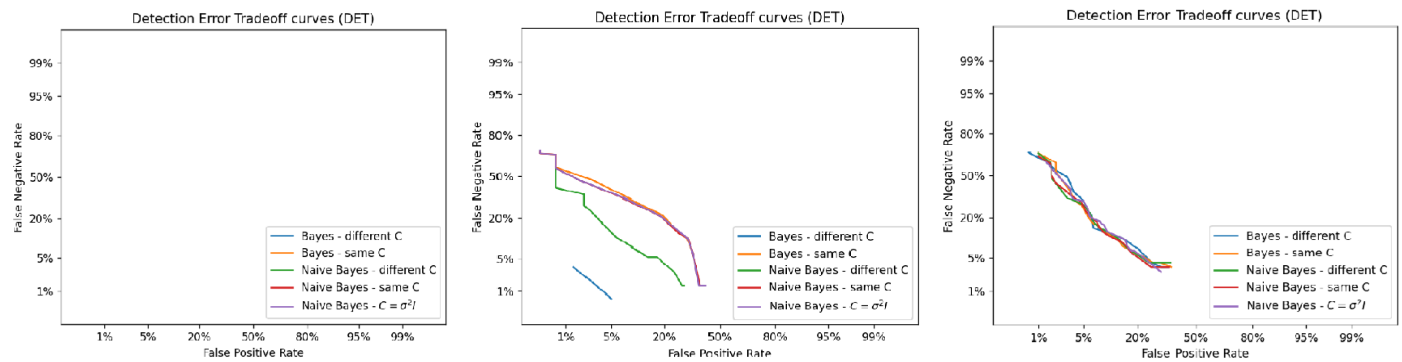

Figure 12: DET curves of all the classifiers for linear data (left), non-linear data (middle), and real-world (right) data

From the decision boundaries given in figures 1 and 3, it is clear that non-linear (quadratic) decision boundaries are obtained when we estimate different covariance matrices for classes, in both the bayesian and naive-bayes classifiers. On the other hand, when we estimate same covariance matrices for the classes, it always results in a linear decision boundary (figures 2, 4, and 5). In the case of linearly seperable data, all the classifiers are able to find the optimal decision boundaries. Because of this, there are no errors reported in the confusion matrices of the classifiers, resulting in an ideal ROC and DET curves for all the classifiers. But in the case of non-linear data, the classifiers that produce a linear decision boundary are not able to find an optimal solution (figures 2, 4, and 5). Hence, the model is said to underfit and we need to improve the complexity of the model. Thus, the classifiers that produce a non-linear decision boundary are able to find a reasonable fit (figures 1 and 3). For real-world data, on visual inspection of the decision boundaries, all the classifiers seem to provide a good split between the instances of different classes.

The above mentioned discussion can be verified quantitatively from the confusion matrices (figures 6 - 10) and from the ROC and DET curves (figures 11 and 12). Larger the area under the ROC curve, better the classifier. It is to be noted that, in the left plot in figure 12, the area under the DET curve for all classifiers is zero. Smaller the area under the DET curve, better the classifier. With these observations, we make the following conclusions.

- **For linearly seperable data, all classifiers perform equally well and they have found the optimal solution**
- **For non-linearly seperable data,**
  Bayes with different Σ > Naive Bayes with different Σ > all other classifiers
- **For real-world data, all the classifiers perform equally well with similar ROC and DET curves**