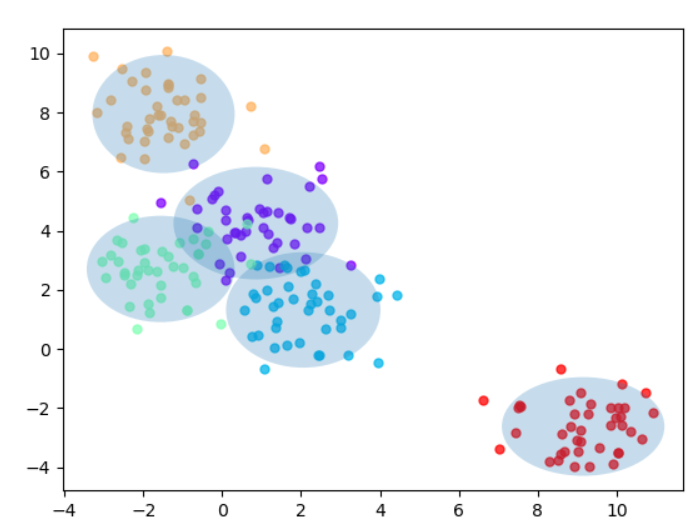
|  |  |
| --- | --- |
|  |  |
|  | Machine Learning |
|  |  |
|  |  |
|  | **Lab 3: Bayesian Learning and Boosting**  LinusGroß  DanielMensah |
|  |  |

|  |
| --- |
|  |

# Bayesian Learning

1. *Maximum-likelihood estimates of μ and ∑*



1. *When can a feature independence assumption be reasonable and when not?*
2. *How does the decision boundary look for the Iris dataset? How could one improve the classification results for this scenario by changing classifier or, alternatively, manipulating the data?*

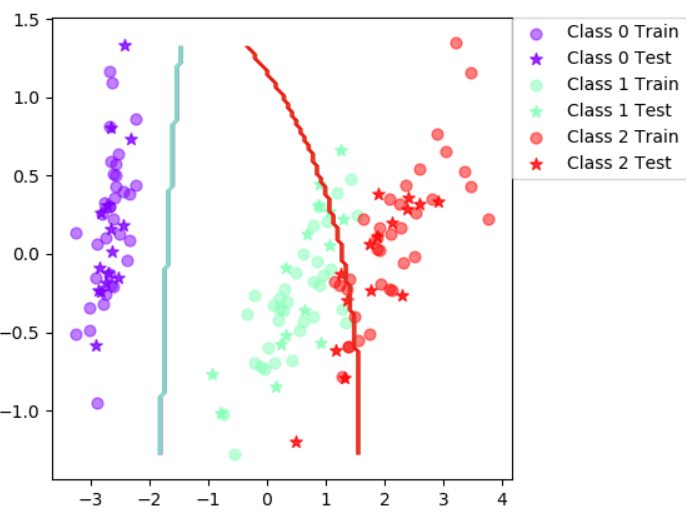


Figure 1: Decision boundary for iris set

The decision boundary doesn’t really fit the boundary between class 1 and class 2. This leads to the misclassification of test and even training samples. This could be due to the fact that these two classes are located close to each other and they even overlap. The decision boundary between class 0 and 1 is however very simple.

A clear decision

# Boosting

*Compute the classification accuracy of the boosted classifier on some data sets using testClassifier from labfuns.py and compare it with those of the basic classifier on the vowels and iris data sets (see Assignment 3):*

1. *Is there any improvement in classification accuracy? Why/why not?*
2. *Plot the decision boundary of the boosted classifier on iris and compare it with that of the basic. What differences do you notice? Is the boundary of the boosted version more complex?*
3. *Can we make up for not using a more advanced model in the basic classifier (e.g. independent features) by using boosting?*

# Decision Tree Classifier

1. *Is there any improvement in classification accuracy? Why/why not?*
2. *Plot the decision boundary of the boosted classifier on iris and compare it with that of the basic. What differences do you notice? Is the boundary of the boosted version more complex?*
3. *Can we make up for not using a more advanced model in the basic classifier (e.g. independent features) by using boosting?*

*If you had to pick a classifier, naive Bayes or a decision tree or the boosted versions of these, which one would you pick? Motivate from the following criteria:*

* *Outliers*
* *Irrelevant inputs: part of the feature space is irrelevant*
* *Predictive power*
* *Mixed types of data: binary, categorical or continuous features, etc.*
* *Scalability: the dimension of the data, D, is large or the number of instances,N, is large, or both.*