

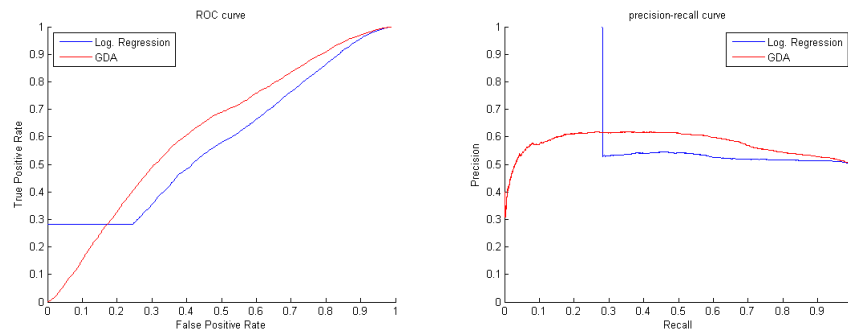
Machine Learning - Homework 2

Michèle Wyss

30.10.2013

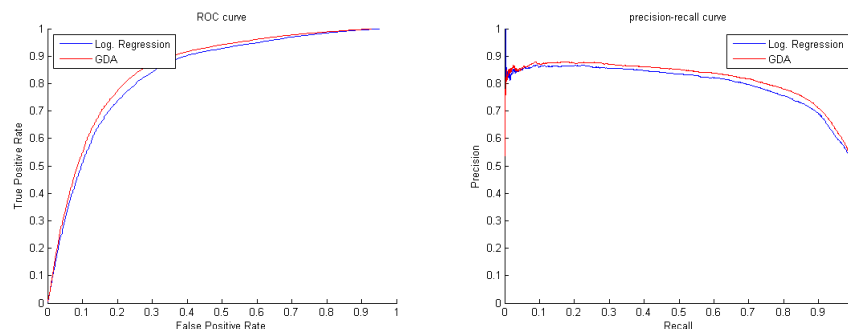
Question 1

Using few training data, e.g. approx. 3% out of all available data, GDA works slightly better compared to the linear regression model. The accuracy is in both cases around 50-65%.

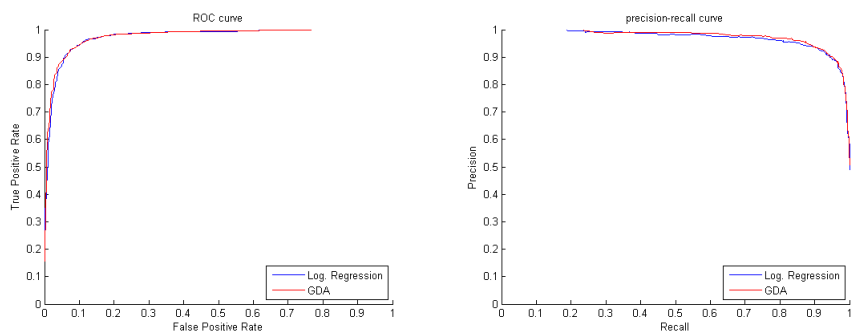


Since GDA does not yield significantly better results than the Logistic Regression Model I'm not sure whether our data is really gaussian. However, because the results using the GDA are more or less stable (mostly between 52-65% correct predictions with different but small training sets), I assume that the data is not very far away from the supposed kind of distribution either.

If we use only slightly more training samples, say 5% (i.e. about 990 images), the results of both the GDA and the Linear Regression model are not very far from each other and both perform rather badly.



Trying out the classifiers with many more training samples leads to very similar results for the GDA as for the Linear Regression model.



So on our problem the GDA model works not bad, but I think that it's still not a satisfying accuracy (since the predictions are correct in only 90% of the test cases).

Question 2