Introduction to Machine Learning - Lab 8

Gustav Sternelöv

Thursday, December 03, 2015

Assignment 1

1.1

1.2

Data is divided into a train and a test data set. The train set consists of 70 % of the original data set and the test set of the remaining 30 %.

1.3

1.4

1.5

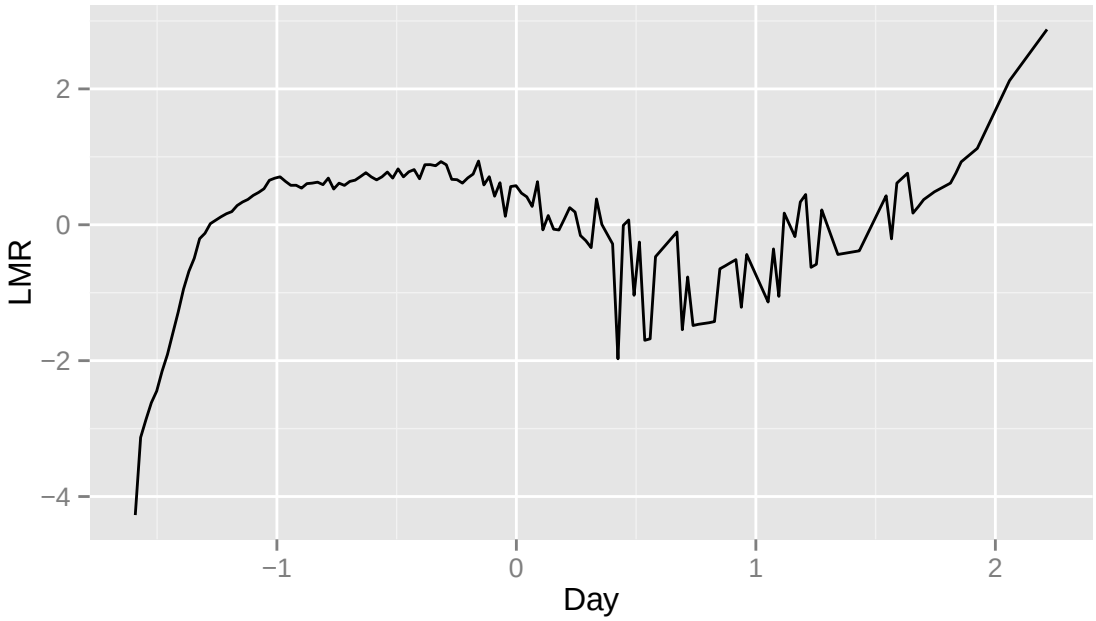
1.6

For the variable Spam the class *-1* is changed to *0*. Then a logistic regression is fitted where Spam is the target variable. The model is fitted on train data and evaluated by predicting the class for the observations in the test set. A confusion matrix of the predicted values against the observed values is created and can be seen below. Since 17 of the 138 observations is misclassified the obtained misclassification rate is 0.123.

```
##          fitted
## test.Spam  0  1
##         0 71 10
##         1  7 50
```

Assignment 2

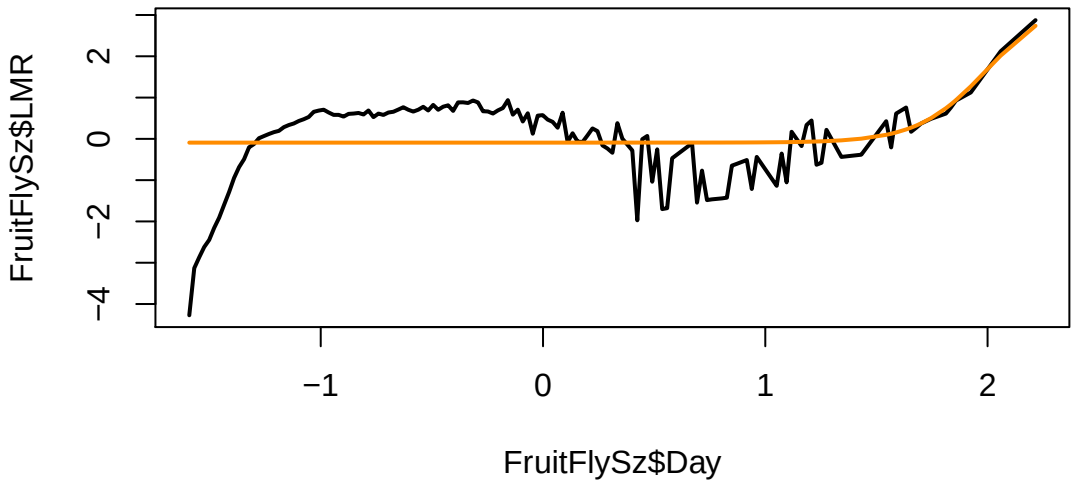
2.1-2.2



2.3

a)

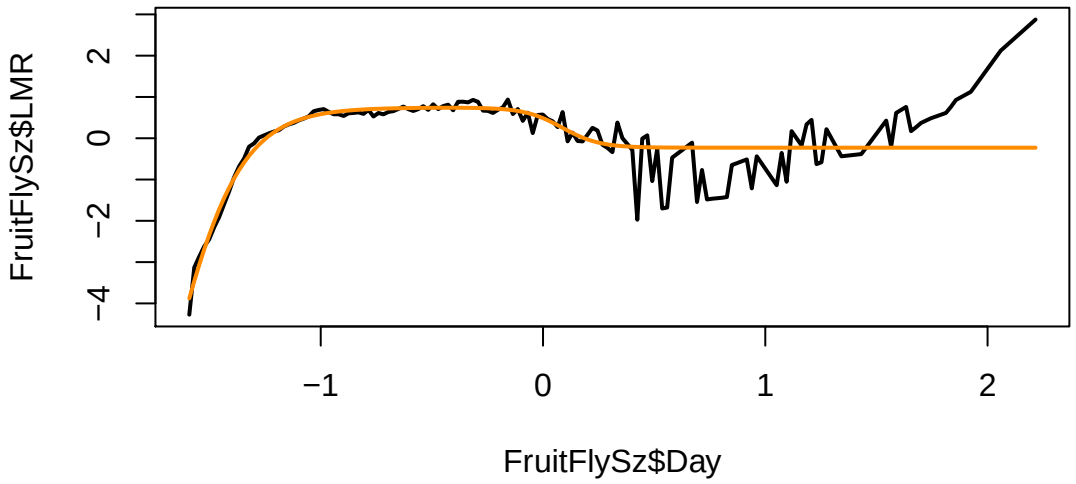
b)



2.4

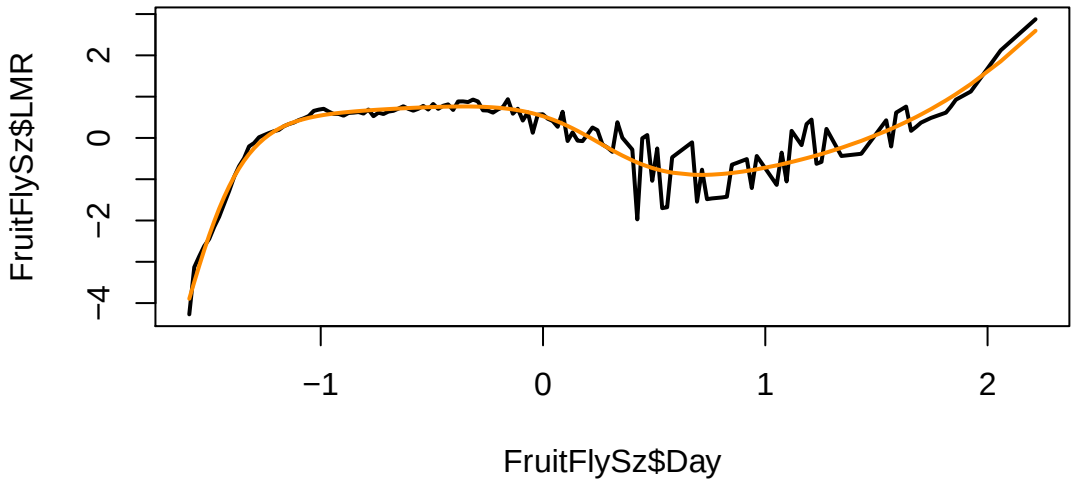
a)

b)



2.5

Four hidden neurons



Five hidden neurons