Homework 5

Brian Lois

2/23/2017

Load the hepatic injury data set using

```
library(AppliedPredictiveModeling)
data(hepatic)
```

See page 8 of the text for a description of the dataset. Ignore the `chem` predictors, and ony use the `bio` predictors.

1. Logistic Regression

For this probelm, do not worry about pre-processing or model tuning.

a. Split the data into a training and test set. Decide whether to use random or stratefied sampling and explain your decision.
b. Use logistic regression and the one-vs-all strategy to build a classification model.
c. Predict the class for each of the samples in the test set.
d. Print the accuracy of your model

2. K-nearest neighbor

Using the same training and test set as above, tune a K-nearest nieghbor alorithm. To do this make a plot of test set accuracy vs k and choose k.

3. Support vector machine

Remove the `Severe` responses from both the training and test set to create a binary classification problem.

Fit several support vector machines to the data. Compare the test set accuracy for different kernal functions and cost parameters. You don't have to do anything fancy here. Just try a few different combinations of parameters and print the test set accuracy.