

## Assignment 4: Validation and Testing

File house-votes-84.txt (attached) contains 435 records of United States representatives. Each record shows how the representative voted on each of 16 different issues (columns in the file).

Based on the voting records, guess the party affiliation of the representative (there are only two parties represented in the United States Congress: Democratic and Republican).

Your tasks:

- 1) Create a classification model of your choice (naive Bayes, logistic regression, Bayesian network, neural networks, etc.)
- 2) Find overall classification accuracy
- 3) Report sensitivity and specificity for each of the two parties
- 4) Measure positive and negative predictive value for each of the two parties
- 5) Plot ROC curve of your model

[http://scikit-learn.org/stable/modules/generated/sklearn.metrics.roc\\_curve.html](http://scikit-learn.org/stable/modules/generated/sklearn.metrics.roc_curve.html)

Please use the five-fold cross-validation method in your tests.