Mid-semester Recap
Classifier I/O

Friday

Supervised Machine Learning as an Adversarial Game

Date: Friday, October 21, 2016
Times: 01:30 PM to 03:00 PM
Presenter: **Dr. Brian Ziebart**
Location: ECS 660

A standard approach to supervised machine learning is to choose the form of a predictor and optimize its parameters based on training data. Approximations of the predictor's performance measure are often required to make the optimization problem tractable. Instead of approximating the performance measure and using the exact training data, this talk introduces a supervised machine learning framework that adversarially approximates the training data and uses the exact performance measure. This formulation provides flexibility for addressing sample selection bias, the fundamental problem hindering advances in active learning, and for inductively optimizing multivariate performance measures like the F-measure and the discounted cumulative gain from information retrieval and ranking tasks.

Recap

- We have come a long way
- You have learned about some of the most popular machine learning techniques
  - Linear Regression (and regularization)
  - Logistic Regression
  - HMMs
  - Naïve Bayes
- Next up: SVMs

Practical Use

- Your labs have given you some hands on experience using these methods
- The lectures have given you the math "under the hood" so that you can use scikit learn properly and effectively

Decision Trees

- When should you use them?
  - For what kind of data?
  - For what kind of applications?
- What about regression trees?
- How can we avoid overfitting?

Linear Regression

- When should you use it?
  - For what kind of data?
- When should you use regularization?

Naïve Bayes

- When should you use it?
  - For what kind of data?
- Why does it work, even when the naïve assumption is violated?

Gradient Descent

- You've also learned about one of the most powerful tools in optimization
- Now, no matter what model you come up with, if its loss/error/likelihood is differentiable, you can find an optimum solution
  - The global optimum if…?

Cross Validation

- You've also learned how to ensure that your models are as strong as you say they are.

Some Challenges

- Lets say we are using logistic regression to do a classification problem
  - What type of features does Log. Reg. expect?
- What can we do if our features are discrete?
  - e.g. color: red, blue, green
  - e.g. Education level: high school, bachelors, master's, PhD

Predict If you will cheat on taxes

- Marital status variable has 4 possible values
  - single
  - married
  - divorced
  - unknown
- We could assign each a number [1,2,3,4]
  - Why might that be a bad idea?

Predict If you will cheat on taxes

- Marital status variable has 4 possible values
  - single
  - married
  - divorced
  - unknown
- We could assign each of the possible values a boolean feature [001,010,100,000]
  - Why is this a better idea?

Multi-class Prediction

- Recall that Naïve Bayes can handle:
  - any number of classes
  - discrete and continuous attributes
    - How?

How can Logistic Regression do multiclass?

- One vs. All
  - E.g. If our possible classes are [A,B,C]
  - Train one classifier for each label
    - A vs not A
    - B vs not B
    - C vs not C
  - Run all 3 on the input
    - How to break ties?
    - What if all classifiers say “not”?
  - Softmax

How can Logistic Regression do multiclass?

- **DAG** approach
  - E.g. **3-class classification**
  - First node is a classifier making the binary decision, **label 1 versus label 3**, say.
  - Depending on the outcome of this decision, the next steps are the decisions **1 versus 2** or **2 versus 3**.