Big Data in Finance - Syllabus

The vast proliferation of data and increasing technological complexities continue to transform the way industries operate and compete. Over the last two years, 90 percent of the data in the world has been created as a result of the creation of 2.5 quintillion bytes of data on a daily basis. Commonly referred to as big data, this rapid growth and storage creates opportunities for collection, processing and analysis of structured and unstructured data.

Financial services, in particular, have widely adopted big data analytics to inform better investment decisions with consistent returns. In conjunction with big data, algorithmic trading uses vast historical data with complex mathematical models to maximize portfolio returns. The continued adoption of big data will inevitably transform the landscape of financial services. However, along with its apparent benefits, significant challenges remain in regards to big data’s ability to capture the mounting volume of data.

The increasing volume of market data poses a big challenge for financial institutions. Along with vast historical data, banking and capital markets need to actively manage ticker data. Likewise, investment banks and asset management firms use voluminous data to make sound investment decisions. Insurance and retirement firms can access past policy and claims information for active risk management.

The course will be a mix of Theory and practice with real big data cases in finance. We will invite guest lecturers mostly for real Big Data Finance Applications.

**Theory**

1. Introduction: What Is Data Science?
2. Structured and Unstructured Data. Providing Structure to Unstructured Data
   1. Classification: k-nearest neighbors, the optimal Bayes classifier, naive Bayes, LDA and QDA, reduced rank LDA, logistic regression. Support Vector Machines.


5. Bayesian Models and Inference using Markov Chain Monte-Carlo (MCMC) including Gibbs sampling and Metropolis-Hastings.


7. Deep learning

8. Estimation and Robustness

9. Optimization Techniques

5. Spam Filters, Naive Bayes, and Wrangling
6. Extracting Meaning from Data
7. Recommendation Engines: Building a User-Facing Data Product at Scale
8. Data Visualization and Fraud Detection
9. Social Networks and Data
10. Data Engineering: MapReduce, Pregel, and Hadoop
11. Big Data in Physics – The CERN case

Practice - Big data in Finance applications

1. Predictive Analytics / Trading
2. Sentiment Analysis
3. Financial Fraud
4. Credit Ratings
5. Pricing
6. Customer Segmentation
7. Know Your Customer
Big Data in Finance - Syllabus

Course Grading

- Homework: 25%
- Final Exam: 35%
- Big data Finance project: 40%

Instructor: Miquel Noguer i Alonso PhD

Textbook

No textbook but recommended reading:

- Doing Data Science, O’Reilly Media
- Principles of Big Data, Elsevier
- Python for Data Analysis, O’Reilly Media
- Machine Learning for Hackers, O’Reilly Media.
- A translation of the R examples in Machine Learning for Hackers to Python can be found here: http://slendrmeans.wordpress.com/will-it-python/
- Probabilistic Programming and Bayesian Methods for Hackers