

## EDUCATION

---

<b>Professional Master's Degree in Statistics</b>	GPA: 3.74/4.00	Graduation Date: Dec 2013
George R. Brown School of Engineering, Rice University, Houston, Texas		
<b>B.A. Degrees in Statistics and Psychology</b>	GPA: 3.41/4.00	Graduation Date: Dec 2012
George R. Brown School of Engineering, Rice University, Houston, Texas		

## SKILLS & QUALIFICATIONS

---

- **Technical Skills:** Python, SQL, R, SAS, Unix Script, Excel, Access, Pig Latin, Tableau, MapReduce
- **Coursework:** Data Science, Data Mining, Quantitative Financial Analytics, Multivariate Analysis, Financial Time Series, Stochastic Process, Risk Management, Regression & Statistical Computing
- **Qualifications:** SAS Certified Programmer for SAS 9

## EXPERIENCES & PROJECTS

---

### **Large-scale Web Graph Processing on AWS** Summer 2013

- Used Hadoop/Pig on Amazon EC2 to analyze a 0.5TB dataset of web graph data (a billion vertices), and computed out-degree histograms of the graph showing the distribution of webpage linkages that demonstrates the connectivity of a web graph
- Studied web link patterns using Page Rank Algorithm for measuring webpages' influences

### **Tweets Sentiment Analysis for Movie Recommendation** Summer 2013

- Derived sentiment scores of real-time tweets and evaluated the public's perception of popular movies based on related tweets using Python and Twitter Streaming API
- Evaluated the similarity between two movies by calculating the correlation based on their sentiment scores and produced a list of recommendations on similar movies

### **Stock Performance and Earnings Visualization on Tableau** Summer 2013

- Used Python to perform analysis on stocks and earning performances of recent IPO companies in technology, pharmaceutical, energy, entertainment and financial industries
- Created an interactive dashboard that demonstrated the Post-IPO analysis by time, industries and filing amount etc. on Tableau for direct communication and easy interpretation of the results

### **ICE Heating Oil Futures Analysis** Spring 2013

- Performed analysis with Excel & R of the monthly trends and seasonality of the daily forward curves for the heating oil futures of the past 3 months, and proposed holistic evaluations for the phenomenon observed
- Predicted general trends of Heating Oil Futures by evaluating the market models and supporting markets

### **Customer Risk Prediction at Nationwide Insurance Company** Summer 2012

Business Analytics Consultant, Marketing & Strategy

- Initiated statistical modeling project that combined psychological theories using SAS & R that assessed customer risk-tolerance levels, to facilitate effective segmentation & targeting of customers
- Acquired proficiency in data retrieval with Teradata database using SQL that ensured concurrency, coherence and quality of analysis, and in-depth understanding of corporate database structure and insurance products
- Delivered methodology & results to senior executives and created sessions to explain methods & tools used to colleagues, aided department in promoting innovative data analysis approaches

### **Kaggle Data Mining Competition** Fall 2012, Spring 2013

- **Titanic Survival Prediction** – Achieved a 90% accuracy in identifying survived passengers by constructing a prediction system using feature scaling and manipulation on Gaussian kernel svm algorithm, integrated with a close inspection of specific trends of the data that overrules the general algorithm
- **Computer Vision** - Constructed a classification system with an ensemble of kernel svm, logistic regression and random forest algorithms to classify 24300 images into five categories and achieved a success rate of 88%

### **Hurricane Damage Modeling at Rice University** Spring 2012

- Modeled frequency of hurricane occurrence and the damage resulted using mixed Poisson regression methods and made predictions, provided basis for risk management in windstorm insurance