
Rice University, Houston, Texas

GPA: 3.80/4.0

Professional Master in Statistics expected May 2014

Area of specialization: Data mining and Financial Statistics

University of Minnesota-Twin Cities, Minneapolis, Minnesota

GPA:3.30/4.0

Bachelor of Science in Economics from 2008 to 2011

SKILLS

Statistical Programming: R Studio/R Language, SAS, Tableau

Database management: SQL, Toad, Vertica

WORK EXPERIENCE

Retail Solutions Inc. (RSi), Research and Development Department, Mountain View, U.S.

Data Science Analyst Intern, Summer 2013

- Joined Leapfrog team aiming next one hundred billion of revenue. Work on project New Product which investigates the consumer goods has the potential of burst sales.
- Acquiring data from database built upon Vertica using SQL query on top of R.
- Cross correlation analysis on census data against retailer report using R and Tableau to discover the impact of population and race and location.
- Focused on new arrival items to illustrate and forecast a common trend of sale lifecycle among different retailers.
- Attended conferences with MaxPoint to present our new service over retail strategy consultant.

RESEARCH EXPERIENCES/PROJECTS

Data Mining: Movie recommender system

11/2012

Using a subset of the Netflix Prize data to build up a recommendation and prediction system which could tell which movie user will enjoy and rate highly. The data consists of 6,040 users who have rated some of the 3,187 movies. Using R implemented multiple models like SVD and KNN to predict the ratings. For the ensemble model blend all the predictors using regression. The final result ranked third among all the competing teams from Rice.

Time Series: Analyzing Global Currency Devaluation Using Gold

10/2012

Using data from 1985 to 2010 to model both value and volatility in order to look at the monthly value how many ounces gold (oz.) 1US Dollar, 1 EU Euro, and 1PRC RMB can buy. Forecasting 1-month to 24-month-ahead prediction and compared it with real values 2010-2012. Used multivariate models to analyze the correlation between those three currencies.

Grappling with Graffiti in San Francisco, CA

10/2012

Analyzed and visualized the available graffiti reporting data for San Francisco, implying the criminal rate of different areas using programming tools R.

Real-time Analysis of 2012 Presidential Campaign Emails

10/2012

Analyzed how the two campaigns react to one another and how they react to real events to predict the pattern of the election periodically using programming tools R.

Analysis of vehicles from three countries

09/2012

Analyzed the impacts of fuel type and the fuel economy on vehicles in order to investigate the market occupancy of vehicles with different fuel type and provided marketing strategy for local dealer using R language.

LANGUAGE: English (Proficient), Mandarin (Native), Cantonese (Native)