Wang Nan (Miya)

(646) 460-5769
Email: nwang26@fordham.edu
Website: https://miyainnyc.github.io/
Github: https://github.com/MiyainNYC

EDUCATION

FORDHAM UNIVERSITY, Information Systems Department **Expected 12/2016**
Master of Science, Business Analytics, GPA: 3.9

- Courses: Web Analytics; Text Analytics; Information Fusion; Big Data; Data Mining for Business; Data Warehousing; Database Management; Business Performance Management and Risk Analytics; Financial Programming and Applications;

NANJING NORMAL UNIVERSITY **2015**
Bachelor of Arts, International Business, Energy Equipment and Environmental Engineering, GPA: 3.6

- Courses: Microeconomics, Global Finance, Visual Basic Programming, Intro to Environmental Engineering, Advanced Mathematics, Fundamental Physics;

EXPERIENCE

Fordham University, Dr. Katsamakas (Area Chair), Dr. Dobin, Information Systems Department **January 2016 – Present**
Research Assistant **New York, NY**

- Research: Quantify and visualize developers' coding commitment and detect their collaborative pattern by conducting social network analysis and time series analysis; (R, Neo4j)
- Research: Quantify picture memorability by training LaMem model; (Convolutional Neural Network on Pycaffe, Ubuntu)
- Assist course teaching of Data Mining for Business by preparing tutorials of algorithmic modeling in Microsoft Azure;
- Crawl historical and streaming tweets and use SQLite for storage; retrieve, transform and analyze unstructured data;(Python)
- Perform academic tutoring of data visualization and business intelligence using Tableau;

Converseon **August 2016 - Present**
Intern, Data Analytics **New York, NY**

- Develop social media metrics to asses online brand performance and customer engagement of a brand and/or a product;
- Generate insights actionable by different stakeholders by retrieving (SQL) and analyzing (sentiment analysis, text analytics) organic and owned social media conversation;
- Help clients (e.g. large pharmaceutical companies and well-known FMCG companies) enhance brand relevancy and customer connection by independently conducting ad hoc research and product analysis;

AXA Hong Kong SL Financial Group **2015**
Intern, Financial Analyst **Hong Kong, China**

- Merge financial data from various sources and conduct time series analysis;
- Predict stock price and Hang Seng Indexby conducting deep-dive qualitative research and building absolute PE model.
- Optimize performance of hedge fund portfolio with help of Markowitz Mean-Variance Model and simulation (R);
- **Top Performer Award** in SL Financial Group in AXA (Hong Kong); "2nd Runner Up" as a team. (Total 20 teams).

Pacific Northwest National Laboratory (PNNL) **2014**
Intern, Business Contracts **Richland, WA**

- Examine contract performance, delivery schedules, and estimates of costs of material, equipment, and production;
- Extract data from company database and paperless system to assist management of multi-million dollar contracts portfolio, acquisition planning and contracting method determination;

PROJECT

Distributed Machine Learning on Google Apps (technically and managerially lead the team)

- Automate data collection and data storage by deploying web crawler on Amazon EC2 and storing streaming text data in S3;
- Achieve faster computational analysis by loading data into Apache Spark framework (Spark SQL), transforming data type into SparseVector (Numpy) and applying principal component analysis (PCA);
- Automate the process of categorizing apps through building a Gaussian Naïve Bayes model (Pyspark MLlib);
- Improve user experience of upgrading apps by summarizing and labeling update descriptions through building a Latent Dirichlet Allocation clustering model Pipeline;
- Help predict app success through analyzing correlation between app download number, app star rate and app category, software version, operating systems, size of the Android Application Package, and etc.,

2016 Fordham March Madness Data Crunch Competition (technically and managerially lead the team)

- Maximize the predictive accuracy of model (RandomForest) through applying techniques such as feature engineering, cross validation, grid searching, hyperparameter tuning and system ensemble; (Sklearn, Pandas and Matplotlib)
- Enhance model interpretability by deploying model as web service (Azure) and developing user-interactive visualization; (R)
- **First Prize** in Fordham March Madness Data Crunch Competition, sponsored by Deloitte;

Others

Programming Language: Python, R, SQL;
Database: MySQL, Amazon Redshift, MongoDB, IBM DB2, IBM Bluemix DashDB
Software: Weka, Databricks, Google BigQuery, Cloudera, SPSS Modeler, Tableau, Spotfire, Qlikview, Cognos, Neo4j;
Certificates: Google Adwords & Analytics; Coursera: Machine Learning, Hadoop Platform and Application Framework;
edX: Excel for Data Analysis & Visualization; Distributed Machine Learning with Apache Spark; Introduction to Linux;
Membership: Steering Community in Fordham Business Analytics Society; Debating Society in Engineering Department in NNU;