

Zhaopeng Liu

M.S in Data Science - Courant Institute of Mathematical Sciences - NYU

• zL1732@nyu.edu • 646-912-3080

SKILLS

Language: Proficient: Python, R, SAS, MySQL, SQL, Matlab **Intermediate:** Tableau, RShiny, C, C++, MapReduce
Software: Jupyter Notebook, Spark(PySpark), Hadoop, Microsoft Excel, MongoDB, Eclipse, Latex

EDUCATION

Master of Science in Data Science, New York University

- Courses: Machine Learning for Data Science, Risk & Portfolio Management, Programming for Data Science(Python), Advanced Python programming, Big Data(Hadoop,Spark)

New York, NY

Graduation: May 2018

Bachelor of Science in Statistics, Minor in CS/Math, University of Illinois at Urbana-Champaign

- Courses: Data Mining, Data Structure, Stochastic Process, Advanced Data Analysis(SAS), Statistical Data Management(SAS), Programming with Java, Categorical Data Analysis
- Dean's List, Graduate with Highest Distinction, Tau Sigma National Honor Society

Champaign, IL

May 2016, Completed

GPA: 3.80

Major: 3.96

WORK EXPERIENCE

7 Chord, LLC: Data Science Intern

- Build unsupervised learning model to cluster over 3000 corporate bonds, trained on AWS
- Worked closely with data engineers and wrote Python scripts that connect with Postgres/MongoDB to prepare and clean the dataset.
- Worked closely with senior trader to develop algo-trading strategy with machine learning

New York, NY

Jul 2017 – In Progress

New York University, Department of Environmental Studies: Research Assistant

- Developed an illegal online wildlife trading detection system with Python
- Build web scraping program to collect information of potential illegal sales and stored them into structured & unstructured data set
- Built classification model with random forest and SVM and achieved 0.9 on the accuracy

New York, NY

Jul 2017 – In Progress

Medici Quantitative Investment: Algorithm Trading Programmer, Intern

- Loaded the real-time options market trading data on tick level and stored them in MongoDB
- Created a prototype of Chinese Futures trading on tick level and implemented it with python, tested the program on copper trading using windows server and profited 2% in two weeks
- Created performance report panel including Shapiro, return indexes for Chinese ETFs market with Shiny on Rstudio to help customers choose their preferred ETFs with a clearer view

Beijing, China

Jun 2016 – Aug 2016

Academic EXPERIENCE

Scikit-Learn Module(Partial) Implementation in Python(In Progress): Long-Term Research

- Implemented hard / soft margin SVM in Pegasos algorithm and tested it on "Spam Email Project"
- Implemented kernelized ridge regression and kernel machine with RBF, Polynomial and linear kernels
- Implemented linear regression based on gradient descent method and stochastic gradient descent method
- Implemented lasso regression based on coordinate descent method

New York, NY

Feb 2017 – May 2017

New York Crime Analysis with PySpark/MapReduce(In Progress): NYU Big Data Course Final Project

- Analyzed New York City crime data set on the computing cluster of New York University
- Wrote PySpark programs to validate data format and to check outliers
- Wrote PySpark programs to discover crime trend over time in different locations
- Developed MapReduce programs to compute the statistics of each variable

New York, NY

Feb 2017 – May 2017

Spam Email Detection System: NYU Data Science Course Final Project

- Parsed raw email data in text format to word tokens using Count Vectorizer in Scikit-Learn module.
- Built models with ML algorithm Naïve Bayes, Decision Tree, Logistic Regression and Random Forest
- Applied penalty factors and created evaluation metrics that focus on minimizing false positive rate
- Performed feature, parameter selection to improve the model, the final model has AUC=0.95 on test data

New York, NY

Oct 2016 – Dec 2016

1st Prize, Contemporary Undergraduate Mathematical Contest in Modeling: Team Leader

- Developed fast dynamic programming algorithm for restoring straight-cut and cross-cut pdf files.
- Won 1st Prize in Beijing, among more than 5000 participants

Beijing, China

Sep 2013 – May 2014