Eric (Binqian) Zeng

40 Memorial Hwy. Apt. 05R, New Rochelle, NY, 10801

+1 929-208-7103 • bz866@nyu.edu; ericzengsearch@gmail.com
https://www.linkedin.com/in/binqian-zeng-257903126/ • https://github.com/bz866

Work Experience

Swiss Reinsurance Company Ltd. (Swiss Re) **Armonk, NY**
Analytics Specialist/Data Scientist *Jan 2019–Present*

∗ **Catastrophe Loss Impact Estimation**
- Estimated liabilities of incoming claims after a catastrophe using satellite images, coverage policies, and building footprints geographic data (geographic zonal statistic analysis, CNN, random forest regression, Python, TensorFlow, ArcGIS)

∗ **Home Renters Insurance Go-to-market Analysis**
- Identified rental insurance market opportunities in New York state using self-designed opportunity scoring metrics(significant test, linear regression); Visualized in an interactive dashboard (Python, JavaScript)

∗ **Insurance Policies Dataset Management**
- Aggregated metadata description for big datasets(billions of records); ETL for data mining(SQL); Backtested hypothesis and correlation(SparkML); Clustering(SparkML)

King Street Capital Management, L.P. **New York, NY**
Data Science Intern (full-time) *Jul 2018–Dec 2018*

∗ **Companies KPI Forecasting Model**
- Forecasted KPIs of companies using an ensemble model; All base and top learners are SVMs with different constraints
- Handled collinearity in meta-features by matrix factorization; Reduced overfitting by ridge regularization and noise injection(SQL, Python, Scikit-learn)

∗ **Analyzing Alpha in Corporate Filings**
- Categorized companies with high and low information ratio based on corporate fillings using RNN and attention mechanism
- Generated sentence representation with word-embedding that is optimized by financial news
- Involved statistical features by capturing textual changes over time(Python, NLTK, TensorFlow)

Crypto Investments **New York, NY**
Machine Learning Engineer Intern *Sep 2017–Dec 2017*

∗ **Event-Driven Forecasting Model for Price of Cryptocurrencies**
- Web-scrapped cryptocurrencies news and trade data by API; Data management for scrapped data(Python, MongoDB)
- News sentiment analysis by Word2Vec; Event embedding by Open IE; Captured effects of different time spans by CNN

Technical Skills & Certificates

- **Programming & Scripting Language:** Python, C/C++, Java, Scala, R, Matlab
- **Toolkits, Softwares & Operating Systems:** Tensorflow, Pytorch, Keras, NLTK, Scikit-learn, AWS, Hadoop, Spark, MySQL, MongoDB, Github, Linux/Unix

Education

New York University, Courant Institute of Mathematical Sciences **New York, NY**
Mathematics in Finance; part-time graduate program on-going *Jan 2019 – Present*
Courseworks: Derivative Securities, Operating System(C/C++), Risk and Portfolio Management with Econometrics, Financial Modeling

New York University, Courant Institute of Mathematical Sciences **New York, NY**
M.S Data Science; GPA: 3.3/4.0; *Sep 2016 – May 2018*
Courseworks: Machine Learning, Natural Language Processing(Kyunghyun Cho), Deep Learning(Yann LeCun), Statistical and Mathematical Methods, Big Data(Hadoop MapReduce, Spark), Data Science in Quantitative Finance, Advance Python Programming

Sun Yat-sen University, School of Engineering **Guangzhou, China**
B.E Theoretical and Applied Mechanics (Fluid Dynamics Focus); GPA: 3.7/4.0 *Sep 2012 – Jun 2016*
Honor: Third-class scholarship (three times)
Courseworks: Numerical Methods, Finite element analysis, Ordinary Differential Equations, Fluid Dynamics

Portfolio

Operating System Components
Keywords: *C/C++ Programming, Operating System*
- Linker; Process Scheduler(discrete event simulation model); Virtual Memory Management Unit; I/O Scheduler

Enhanced Seq2Seq Model for News Text Summarization (Capstone Project in NYU)
Keywords: *Python, Pytorch, Bidirectional-LSTM, Attention Mechanism, Pointer Network, Semantic Relevance*
- The Seq2seq text summarization model is a hybrid of extractive and abstractive approaches trained on CNN/Daily Mail news
- Bi-LSTM encoder with attention mechanism; Pointer network improves the accuracy of words generation and the ability of handling out-of-vocabulary words
- Involved cosine similarity term between encoded source representation and generated summarization representation into the negative log-likelihood loss function to encourage semantic relevance

Commodity Trading Position Forecasting Model
Keywords: *Python, Alpha Strategy, Matrix Factorization, Regularization, Linear Regression*
- Created a regression model that forecasts the optimal trading positions for oil, sugar, copper, gold, and natural gas futures
- Estimated covariance models; Explained the variance asset class by top alpha factors; Signal filtering down by SVD