Seung Jae Bang

365 Bond St. Apt C513, Brooklyn, NY. 11231.
(347) 325-0765 · sb2927@columbia.edu

PROFESSIONAL EXPERIENCE

Kensho Technologies / S&P Global — New York, NY
Data Scientist — Mar 2018 – Present

- Productionized market sentiment analysis model in Python, to predict company level sentiment on news corpus daily based on bag-of-words approach. Developed target-dependent sentiment model by using LSTMs to capture context information around the target company, improving the sentiment granularity.
- Research / prototyped methods to summarize company earnings transcripts based on extractive summarization, highlighting important analyst questions based on key phrases.
- Prototyped implementation of weak-supervision, to expand news sentiment labels while reducing human effort, based on Snorkel.

Goldman Sachs — New York, NY
Vice President, Quantitative Analyst (Interest Rate and Mortgage products) — June 2010 – Feb 2018

- Designed and productionized risk models based on statistical methods, which include:

- Designed risk-metric forecasting model using Principal Component Analysis on macro factors, to test the capital adequacy of the firm's trading desk.
- Implemented P&L decomposition (regression based) model for interest rate products by identifying the significant risk factors, and attributing the drivers of P&L to risk factors.
- Designed missing data imputation method for sparse (non-daily) Home Price Index time series, using regression techniques that incorporated autocorrelative and seasonality effects, to be used for daily risk factor simulation.
- Designed data imputation techniques to backfill time series data, through the use of Gaussian Mixture Model.

Associate, Market Risk Analysis (FX and Mortgage products) — Apr 2010 – June 2014

- Analyzed risk models in conjunction with market activity to explain risk changes to market factors.

Prudential Financial — Newark, NJ
Part Time Intern, Asset Management — Sept – Dec 2009

- Automated the process of rebalancing 150 strategic portfolios of asset classes for retail investors based on their risk/return appetites, using portfolio optimization techniques.

RESEARCH PROJECTS

Inferring Cultural Fit in Organizations from Language [Code Link]
Advisor: Sandra Matz, Columbia University — Aug 2019 – Present

- Processed large amount of Linkedin profiles (~60million) to extract personal statement text
- Analyzed linguistic style using LIWC (Linguistic Inquiry and Word Count) and study the relationship between employee terms of stay and their deviation from the organization's linguistic style (in progress)

Predicting Stock Market Behavior from Social Media [Code Link]
Advisor: Joseph Johnson, University of Connecticut — May 2019 – Present

- Implemented sentiment analysis classifier using StockTwits data and analyzed the relationship between sentiment and stock returns (in progress)

OTHER PROJECTS

Tumor Image Detection on Camelyon 16 Lymph Node Images [Code Link]
Course final project, Columbia University — Oct – Dec 2018

- Implemented image classification architecture combining transfer learning with multiple inputs to incorporate tissue images at multiple zoom levels simultaneously.

Kaggle Airbus Ship Detection Challenge [leaderboard]

- Implemented image segmentation framework (U-Net) in Keras
- Our team (of 8 people) achieved top 1% (9^{th} among 882 teams)

EDUCATION

Columbia University, School of Engineering and Applied Science — New York, NY
M.S. in Data Science (part-time enrollment), *GPA: 3.93 / 4.00* — Sept 2016 - Present

Columbia University, School of Engineering and Applied Science — New York, NY
M.S. in Financial Engineering, *GPA: 3.93 / 4.00* — July 2008 - Dec 2009

Cornell University, College of Engineering — Ithaca, NY
B.S. in Electrical and Computer Engineering, *GPA: 3.82 / 4.00*, ***Magna Cum Laude*** — May 2008

AWARDS

Korean-American Scientists and Engineers Association Scholarship — Aug 2007
John McMullen Dean's Scholarship in Engineering for Academic Excellence — Aug 2004 - May 2008

ACTIVITIES & TEST

Lead Instructor at Data Science Academy, *S&P Global* — Apr - Sept 2019

- Company initiative to train in-house employees in data science
- Instructed generalized linear models, ANOVA, data visualization

New GRE: Verbal (166 / 97%), Quantitative (170 / 96%), Writing (4.0 / 57%) — July 2019

SKILLS

Programming	Python, Matlab, SQL
Deep Learning Framework	Tensorflow, PyTorch
Operating System	Linux, Windows