Taehee **Jung**

Statistics Master student @UCB & former data analyst @KCB for 5+ years.

5424 Fifth Ave #106, Pittsburgh, PA 15232

412-403-8832 | taehee_jung@berkeley.edu | http://theejung.github.io

Summary

I am interested in analyzing customer/industrial data and applying recent deep/machine learning techniques to them. Especially, I am interested in recent neural networks (e.g., LSTM, GAN) and machine learning techniques on data mining and natural language processing applications. Currently, I am looking for a **full-time/internship** positions in **data/research scientist**.

Education

University of California, Berkeley — *Berkeley, CA, USA*

Master of Arts, Department of Statistics — *Aug 2016 - May 2017*

- Selected courses: Machine Learning (A), Linear Models (A), Advanced Probability, Statistics, and Statistical Computing (A)

Korea Advanced Institute of Science and Technology — *Daejeon, Korea*

Bachelor of Arts, Department of Management Science — *2007 - 2011*

- Academic excellence scholarship (2007 - 2011)

Daeil Foreign Language High School — *Seoul, Korea*

Experience

University of California, Berkeley — *Berkeley, CA, USA*

Empirical Examination on Generative Adversarial Models, ML class project [report] — *Feb-May 2017*

- Studied vanilla Generative Adversarial Networks (GAN), deep convolutional GAN, Wassertein GAN, and Cycle GAN
- Compared different GANs on generation, interpolation, projection, arithmetic operations, translation, and completion tasks.
- Received full credit among ML projects, and extended the empirical study to Kinship verification research

S&P 500 Stock Portfolio with Sentiment Analysis, Capstone Project [report] [ppt] — *Jan - May 2017*

- Analyzed sentiment of tweets predicting volatility and developed autoregressive predictive models
- Constructed a stock portfolio using Markowitz objective function with l1 penalty
- Extended the sentiment analysis to CSpikes research for publication

Creating R package for MFA, class project of Statistical Computing [demo] — *Nov - Dec 2016*

- Developed R package for Multiple Factor Analysis to perform mfa, bootstrap, and plot the results and visualized the results on Shiny app

Korea Credit Bureau (KCB) — *Seoul, Korea*

Big, Public Data Analysis, funded by Ministry of Government Administration & Jeonju City — *Sep - Dec 2015*

- Analyzed and visualized boarding histories, local floating population, petition documents, and home to office O/D information
- Constructed new strategies for bus rearrangement and adopted by Jeonju city on Feb, 2017
- Practically chosen to Jeonju city's bus rearrangement plan [article] [report] (written in Korean)

Fraud Detection System Development, funded by Woori Bank — *Dec 2014 - Aug 2015*

- Segmented customers' transaction histories and developed predictive ensemble models on regression, neural network, and decision tree
- Compared with previous FDS, overall detection rate increases more than 20% [article] (written in Korean)

Migrant Analysis for Magok New Town, funded by LG Economic Research Institute — *Jun - Sep 2014*

- Developed demographic database of migrants and classified them using decision tree for potential cultural demand to the new town

Location based Decision Support System, funded by Korea Credit Bureau & R114 — *May - Aug 2012*

- Developed an integrated database (e.g., location, real estate, residential's demographic and financial data) called K-atlas

Publication

CVPR'18 **Kinship Verification with Facial Feature Heredity using Attentive CycleGAN**, Taehee Jung*, Dongyeop Kang*, In Preparation, preprint

AAAI'18 **CSpikes: Detecting Causality between Time Series and Textual Data**, Dongyeop Kang, Taehee Jung, Ang Lu, Zheng Chen, Eduard Hovy, Submitted, preprint

Skills

Programming Python, R, SAS, SQL, Unix/Linux, Tableau, QGIS

Language English (Fluent), Korean (Native), Spanish (Beginner)