Seungjun "Josh" Kim

76000 Verano Road, Irvine, CA 92617 | kims17@uci.edu

Website: seungjun-data-science.github.io | LinkedIn : /kimseungjun | Github: /seungjun-data-science

EDUCATION

University of California, Irvine   Irvine, CA | June 2028 (Expected)

Doctor of Philosophy (Ph.D) in *Informatics*   GPA: N/A

Advisor: Sean D. Young

Vassar College   Poughkeepsie, NY | May 2020

Bachelor of Arts (B.A.) in *Mathematics (with Departmental Honors);* Minor in *Computer Science*   GPA: 3.80 / 4.0

- **Coursework**: Applied Statistical Modeling, Data Science and Statistical Learning, Computational Linguistics, Bayesian Inference, Real Analysis, Econometrics, Linear Algebra, Analysis of Algorithms

Oxford University, St. Catherine's College   Oxford, United Kingdom | Jan – Jun 2019

Junior Year Abroad Visiting Students Program

- **Coursework**: Statistical Machine Learning, Quantitative Economics, Finance, Corporate Social Responsibility

RESEARCH INTEREST

Broadly interested in:

- Data Science for Social Good and Health Informatics.
- Addressing burnout of healthcare workers, in particular, nurses.
- Pre-emptively detecting spread of virus or disease, severe depression, suicide risks, gun violence, drug overdose/addiction using unstructured data such as social media feeds.
- Developing new methodologies to better detect misinformation in the healthcare domain and understand how they affect behaviors, perceptions and responses of patients or the general public towards diseases or public health policies.
- Identifying bias in health-tech systems or Machine Learning models used in healthcare and present ways to make them more equitable, transparent, and fair.

RESEARCH EXPERIENCE

Institute for Social Research, University of Michigan, *Research Area Specialist*   Ann Arbor, MI | July 2020 – June 2022

- The purpose of the research with the CDC was to measure the cost of domestic violence in the U.S.
- Wrote scripts in STATA and Python to:
  - Identify perpetrators of domestic violence based on offense description from the Criminal Justice Administrative Records System (CJARS) database.
  - Link these offenders with family relations and residence data from the U.S. Census Bureau and implemented logics to identify victims of intimate partner violence (IPV).

Vassar College Economics Department, *Research Assistant*   Poughkeepsie, NY | May – Dec 2018

- Researched with Dr. Esteban Argudo on two topics:
  - Describing the expected impact of the 2017 Repatriation Tax Holiday on the U.S. economy using 10-K data collected from the Securities and Exchange Commission for fiscal years 09'-17' on 10,148 unique companies.
  - Measuring the impact of macroeconomic effects on loans using loan data from All Lending Club.

WORK EXPERIENCE

Institute for Social Research, University of Michigan, *Research Area Specialist*   Ann Arbor, MI | July 2020 – June 2022
- Harmonized criminal justice data to a schema specified by the Criminal Justice Administrative Records System (CJARS).
- Built an automated and schedulable Python program for data quality assurance that applies anomaly detection algorithms and business logic to detect errors and missing-ness in the data.
- Wrote SQL queries that produces statistics as part of the Django program that generates statistical reports from the CJARS database for agencies such as the Department of Corrections.
- Performed literature review to be included in grant proposals for institutions including the National Institute of Justice (NIJ) and the National Science Foundation (NSF).

Spotify, Mobile Web Player Team, *Data Science Intern*   New York, NY | Jun – Aug 2019
- Performed user cohort analysis in Python on user data to better understand user behavior and needs.
- Used anomaly detection techniques to identify deep link error which affected, on average, 14.8% of daily active users.
- Predicted growth projections of new features or marketing methods in collaboration with the UX Researcher and Product Manager to make proposals for roll out.

Family Services Inc., *Data Analytics Intern*   Poughkeepsie, NY | Jan 2018 – May 2018
- Cleaned employee data and investigated factors associated with employee turnover using Pearson correlation, partial dependence plots, SHAP values, permutation importance and LIME.
- Created an ensemble model that averaged the predicted probability of turnover from three solo models, Logistic Regression, Random Forest and Adaptive Boost, and classified employees into different turnover risk bins.
- Human Resources Manager used this model to preemptively address dissatisfaction of employees who are predicted to have high likelihood of quitting the job.

SKILLS

Data Analysis, ML, DB, Visualization: Python, Tensorflow, Pytorch, R, STATA, Tableau, MySQL, BigQuery, HTML, CSS

Other: Git, LaTex, Asana, Microsoft Suite, Google Suite, Prezi, Canva

Language: Korean (native), English (native), Chinese (intermediate; HSK Level 5 Certificate)

COMPETITIONS, HONORS, AWARDS

Vassar College DataFest, *Best Overall Analysis Award*   2020

Kaggle, *Instant Gratification*   2019
- Binary classification competition with synthetic data provided by Kaggle. Ranked 328 out of 1832 (Top 18%).

Kaggle, *Santander Customer Transaction Prediction*   2019
- Binary classification competition using Santander's customer transaction data to predict which customers will make a specific transaction in the future. Ranked 1354 out of 8802 (Top 16%).

Vassar College, *Tananbaum Fellowship*   2019
- Awarded to 20 junior year students selected based on merit.
- Provided $5,000 worth of financial support for internships and career guidance.

National Model United Nations (NMUN) NY, *Outstanding Delegation & Position Paper Awards*   2015
- Received the Outstanding Delegation Award as a member of the delegation (Hudson River Group - Dutchess Community College and Vassar College Joint Team) representing Australia.
- Received the Outstanding Position Paper Award at the Nuclear Non-Proliferation Treaty Committee.

LEADERSHIP

Vassar Haiti Project (VHP), *Student Leader and Treasurer*   Poughkeepsie, NY | Jun 2019 – May 2020
- Led general board members as one of the student leaders at the VHP, a 501(c) non-profit supporting the welfare of artists and sustainable development in Chermaître, Haiti through auction and sale of Haitian art and handcrafts.
- Allocated budget to events, digitized receipts on QuickBooks, managed money flow, worked as cashier during fundraiser art sales using the Square app, and educated new treasurers.

Korean Defense Intelligence Command, *Army Sergeant and Squad Leader*   Anyang, South Korea | Aug 2015 – May 2017
- Led 35 squad members during military training sessions as squad leader.
- Translated open source intelligence (e.g. editorials and think-tank reports) regarding military, national security, politics, and international relations of North Korea and the U.S. from English to Korean and vice versa.
- Implemented a peer review system for the translations.

COMMUNITY SERVICE

Liberty in North Korea (LINK) Vassar Chapter, *Volunteer*   Poughkeepsie, NY | Sep 2014 – May 2015
- Raised $2,000+ funds for rescuing North Korean refugees via the Asian Night Market sales and the Benefit Concert.
- Increased awareness on the humanitarian crisis in North Korea by inviting North Korean refugees for talks and screening documentaries or movies on human rights abuses in North Korea.

Gwan-Ak Community Center, *Volunteer*   Seoul, South Korea | Jul 2007 – Aug 2013
- Delivered lunch boxes for the elderly living alone who are unable to visit the community center for meals.
- Visited homes of the elderly on a regular basis for counseling.
- Initiated the *Hand-Written Letter Program* to enhance emotional health of the elderly which matched the senior citizens living alone with volunteers who sent hand written letters to them.