

Shaun (Siyeon) Kim

(434) 227-3351 | sk981102@gmail.com | [Personal Website](#)

Education

Columbia University, New York City, NY

Expected December 2022

M.A. in Statistics / Data Science

University of Virginia, Charlottesville, VA

August 2017 – May 2021

B.A. in Statistics and Computer Science, graduated with Distinction

- **GPA:** 3.73/4.0 | **Major GPA:** 3.8/4.0 | **Relevant Coursework:** Cloud Computing, Artificial Intelligence, Probability, Data Structure, Sampling Methods, Machine Learning, Algorithm, Database, Time Series

The Loomis Chaffee School, Graduated Cum Laude

Sept 2014 - May 2017

Work Experience

Gentine Lab | Columbia University

June 2020 - Present

Graduate Researcher

- Worked closely with the professor and post-doctorates to develop a machine learning model such as XG Boost, CNN, and LSTM that accurately predicts carbon dioxide flux using Python
- Pre-processed and created data visualization for 5-dimensional geographical data using xarray and hvplot

Clova AI | Naver Corporation

June 2020 – Aug 2020

Machine Learning Engineer Intern

- Developed and fine-tuned a toxic reply filtering model for NAVER news forum by utilizing BERT transformer, which resulted in an F1 score of 94.87 and had an experience of creating a baseline model using LSTM and RNN
- Developed an algorithm that will select topics from a news article utilizing search trend spike, built a demo using JavaScript to demonstrate the algorithm, and familiarized other Natural Language Processing tasks in doing so
- Utilized Spark to fetch and build a data pipeline from Hadoop and preprocessed the data

Teaching Assistant for Machine Learning

Jan 2020 – May 2021

Teaching Assistant

- Worked closely with the professor to develop the course material for Introduction to Machine Learning (CS 4774) and evaluated the performance of the students
- Provided help to students during office hour with various machine learning algorithms and applications

Boichuk Lab | McIntire School of Commerce

Aug 2019 – May 2020

Undergraduate Data Science Researcher

- Conducted data analysis using R on benefit corporations and the correlations between the companies and interests from the public by utilizing different regression techniques
- Designed research methods using literature reviews, designing surveys, collecting raw data from scraping Google trend data

Darden School of Business

April 2019 – Aug 2019

Product Management Intern

- Built a new, leading-edge website on Drupal 8 platform by collaborating with web-developers, designers, communication and marketing team, and individual departments of information
- Determined design compositions, site maps, and image qualifications for over 800 webpages based on the user feedback, website traffic and A/B testing result from the previous Darden website

Projects

Crowd-source Data Platform

Sep 2020 – Nov 2020

- Led a team of four through application design and development, which won the Social Innovation scholarship at Yonsei University
- Built a web application that allows people to crowd-source data collection; admin can request specific data collection tasks, and the regular users can submit data and evaluate data quality

Facebook Data Challenge

March 2019

- Presented a data-driven pitch, which analyzed the business landscape of San Francisco, utilizing Tableau and R
- Selected as a finalist for the Data Challenge NYC, a selective, two-day career development program for data scientists
- Was offered an internship as a Data Scientist, which was canceled due to Covid-19

Skills & Interests

Technical: Python (Proficient), R(Proficient), Java (Proficient), MySQL(Proficient), Tableau (Intermediate), JavaScript (Intermediate), C++ (Intermediate), SAS(Elementary), Spark (Elementary), Hadoop (Elementary), Microsoft Suite

Languages: Fluent in English and Korean, Intermediate proficiency in Chinese and German

Other Studies and Projects: Relations Between Air Pollution and Happiness, UVA Wheelchair Access, Benefit Corporation Crawling + Analysis, Toxic Reply Modeling, News Topic Modeling, Covid-19 Flight Visualization, DebunkVax