

SANGJUN (SAMUEL) Ko

Urbana, Illinois 61801 | sangjun2@illinois.edu | www.linkedin.com/in/sangjunko

EDUCATION

University of Illinois at Urbana-Champaign

Bachelor of Science in Statistics (CGPA: 3.83/4.00)

Champaign, Illinois

January 2023 - May 2025 (Expected)

- Mathematics & Data Science Minor

University of Michigan-Shanghai Jiao Tong University Joint Institute

Shanghai, China

Bachelor of Engineering in Mechanical Engineering (CGPA: 3.50/4.00)

September 2019 - December 2020

- First Class Scholarship

WORK EXPERIENCE

Applied Technologies for Learning in Arts & Sciences

Champaign, Illinois

AI/ML Intern Team Leader

January 2024 - Present

- Lead a team of 5 interns.
- Develop predictive machine learning models with large data sets(100,000+ entries) using Python
- Utilize advanced techniques: XGBoost, Random Forests, Regression, Neural Networks, BERT, Transformers
- Collaborate closely with 3 coordinators in weekly meetings to align project goals and track progress

STAT107 Data Science Discovery

Champaign, Illinois

Course Assistant

August 2023 - Present

- Facilitate learning for over 1000 students by assisting students in grasping fundamental data science concepts by conducting interactive sessions, and providing clarifications on course materials
- Manage Git for learning platform modifications, benefiting 1000+ students; create, edit, and proofread course materials
- Collaborate with a team of 100+ staff and 2 professors in weekly meetings

KM Internacional S.A.

Managua, Nicaragua

Intern

May 2023 - August 2023

- Oversaw comprehensive inventory management and streamlined administrative tasks for a small textile company
- Lead a team of three workers to achieve a 98% accuracy rate in record-keeping and a 15% reduction in operational costs

PROJECTS/ AWARDS

3rd Place, 2024 COVID Information Commons Student Paper Challenge

September 2024

- Analyzed the socioeconomic and mental health impacts of COVID-19 in Mexico, identifying that 42% of females and 40% of young adults faced job loss.
- Proposed data-driven interventions to support vulnerable groups through geospatial and statistical analysis.

Paraphrase Detection

July 2024

- Achieved 90%+ accuracy in paraphrase detection using BERT, RoBERTa, and SBERT models on the MRPC dataset
- Developed a Gradio interface for real-time paraphrase detection, facilitating easier evaluation of textual similarities for academic and business use cases.

International Student Forecasting

April 2024

- Processed and analyzed large-scale student enrollment data (over 100,000 entries) from UIUC to identify trends and patterns
- Suggested resource allocation optimization by forecasting international student numbers using linear regression and deterministic processes

SKILLS

Languages: English (Fluent), Spanish (Fluent), Korean (Fluent), French (Basic), Chinese (Basic)

Computer: Python (Advanced): Pandas, NumPy, Scikit-learn, TensorFlow, NLTK, Matplotlib, Seaborn, BERT, Random Forest, XGBoost, Gradio
R (Intermediate): Tidyverse, ggplot2, dplyr, caret
Git (Intermediate), SAS (Basic), Tableau (Basic), SQL (Basic), Office (Advanced)

Certificates: Data Science DISCOVERY, Data Science Certificate, Coursera Business Analytics Certificate

COMMUNITY SERVICE AND EXTRACURRICULAR ACTIVITIES

Treasurer in Performance UIUC Touch

February 2023 - May 2024

- Strategically plan and execute 10 fundraising initiatives to support team activities, costumes, and travel expenses, resulting in a 30% increase in available funds compared to the previous year