

# SANGJUN (SAMUEL) Ko

202 E White St, Champaign, Illinois, 61820 | 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)
  - Mathematics & Data Science Minor
- Champaign, Illinois  
January 2023 - May 2025 (Expected)
- University of Michigan-Shanghai Jiao Tong University Joint Institute**  
Bachelor of Engineering in Mechanical Engineering (CGPA: 3.50/4.00)
  - First Class Scholarship
- Shanghai, China  
September 2019 - December 2020

## WORK EXPERIENCE

- Applied Technologies for Learning in Arts & Sciences**  
*AI/ML Intern Team Leader*
  - 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
- Champaign, Illinois  
January 2024 - Present
- STAT107 Data Science Discovery**  
*Course Assistant*
  - 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
- Champaign, Illinois  
August 2023 - Present
- KM Internacional S.A.**  
*Intern*
  - 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
- Managua, Nicaragua  
May 2023 - August 2023

## PROJECTS/AWARDS

- 3rd Place, 2024 COVID Information Commons Student Paper Challenge**
  - 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
- September 2024
- Paraphrase Detection**
  - 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
- July 2024
- International Student Forecasting**
  - 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
- April 2024

## 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

- Consultant in Illinois Business Consulting**
- September 2024 - Present
- Collaborating with interdisciplinary teams to deliver strategic solutions for real-world business challenges, including market analysis, operational optimization, and growth strategy, for clients ranging from Fortune 100 companies to nonprofits