

JUNWON CHOI

☎ (310) 254-4579 ✉ cjunwon.tech@gmail.com 🔗 linkedin.com/in/cjunwon 🌐 cjunwon.github.io

EDUCATION

University of California, Davis <i>Master of Science, Statistics (Data Science Track)</i>	Davis, CA <i>December 2025</i>
University of California, Los Angeles (UCLA) <i>Bachelor of Science, Data Theory</i> <ul style="list-style-type: none">• GPA: 3.7 / 4.0• Relevant Coursework: Machine Learning, Generative Data Science, Mathematical Modeling, Spatial Statistics	Los Angeles, CA <i>June 2024</i>

EXPERIENCE

Sensing and Robotics for Infrastructure (SRI) Lab at UCLA <i>Research Intern</i> <ul style="list-style-type: none">• Led multiple projects to add key features on a robust road assessment web application funded by the City of Los Angeles• Integrated parallel processing scripts to AWS S3, EC2 pipeline, automatically restructuring 10+ GB of geospatial data• Evaluated road usage patterns and traffic congestion in a network of 88+ million node pairings using network analysis• Customized computer vision models to classify parked cars from 150+ GB of onboard camera footage	Los Angeles, CA <i>March 2023 – Present</i>
The Daily Bruin <i>Data Team Director</i> <ul style="list-style-type: none">• Directed a team of 20 journalists to publish 50+ articles and independent data visualizations in the newspaper• Launched two additional production teams and trained 15+ new reporters, expanding production capacity threefold• Reported GA4 engagement data for all 17 internal teams and advised content strategies to increase readership <i>Data Journalist</i> <ul style="list-style-type: none">• Pitched and produced data-driven, interactive, and responsive web articles using React, JavaScript, and HTML/CSS• Interviewed and liaised with external organizations and individuals for dataset requests and story details	Los Angeles, CA <i>May 2023 – June 2024</i> <i>October 2022 – May 2023</i>
DataRes at UCLA <i>Research Team Lead</i> <ul style="list-style-type: none">• Headed a team of 5+ students to develop machine learning projects based on novel geometric deep learning research• Designed custom SQL server to host and filter 50+ GB of raw data, increasing data distribution efficiency by 93%	Los Angeles, CA <i>October 2022 – March 2024</i>
Republic of Korea Marine Corps <i>Sergeant (Communications Squad Leader)</i> <ul style="list-style-type: none">• Supervised and trained a 20-person team to set up radio and wired communication networks for battalion operations• Interpreted and translated Korean-English for ROK and U.S. officers during 3 major joint training operations	Gimpo, South Korea <i>October 2020 – April 2022</i>

PROJECTS/RESEARCH

Synthetic Survey Data Generation and Evaluation: A Case Study on National Voter Files <ul style="list-style-type: none">• Built end-to-end pipeline to generate synthetic data using U.S. voter files with parametric, GAN, LLM based models• Evaluated synthetic data quality using 4 general utility, 1 target-specific utility, and 5 privacy metrics
Fingerhut (Online Retailer) Strategic Consulting [<i>Capstone Project – Data Theory, B.S.</i>] <ul style="list-style-type: none">• Analyzed 1.6+ million customers' web sessions to extract key features driving account activation and purchases• Developed custom stochastic, classification, and neural network models to visualize and predict customer behavior• Presented findings to Fingerhut executives and recommended strategic changes to increase customer retention
Research Paper Classification Optimization Using Networks and NLP <ul style="list-style-type: none">• Implemented k-means and SVD to simplify categorization of 2.4+ million arXiv research papers• Optimized KNN, random forest, and fastText to develop a high accuracy paper classification model• Integrated ANNOY (Approximate Nearest Neighbors) model to retrieve similar papers within 800ms

SKILLS

Programming:	Python, R, SQL (MySQL, SQLite), JavaScript, HTML/CSS, L ^A T _E X
Data Visualization:	Tableau, Chart.js, D3.js, matplotlib, seaborn, ggplot2, Plotly, QGIS
Machine Learning/Networks:	PyTorch, TensorFlow, scikit-learn, tidymodels, NetworkX, igraph
Development/Analytics:	Git, AWS (S3, EC2, RDS, Lambda), Databricks, Google Analytics, Excel
Languages:	English and Korean (bilingual fluency)