

# JUNWON CHOI

☎ (310) 254-4579 ✉ [cjunwon.tech@gmail.com](mailto:cjunwon.tech@gmail.com) 🔗 [linkedin.com/in/cjunwon](https://www.linkedin.com/in/cjunwon) 🌐 [cjunwon.github.io](https://cjunwon.github.io)

## EDUCATION

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

## EXPERIENCE

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

## PROJECTS/RESEARCH

<b>Synthetic Survey Data Generation and Evaluation</b> [ <i>ACM KDD 2025 - ADS Track</i> (submitted, under review)]
<ul style="list-style-type: none"><li>• Built end-to-end pipeline to generate synthetic data using U.S. voter files with parametric, GAN, LLM based models</li><li>• Evaluated synthetic data quality using 4 general utility, 1 target-specific utility, and 5 privacy metrics</li></ul>
<b>Fingerhut (Online Retailer) Strategic Consulting</b> [ <i>Capstone Project – Data Theory, B.S.</i> ]
<ul style="list-style-type: none"><li>• Analyzed 1.6+ million customers' web sessions to extract key features driving account activation and purchases</li><li>• Developed custom stochastic, classification, and neural network models to visualize and predict customer behavior</li><li>• Presented findings to Fingerhut executives and recommended strategic changes to increase customer retention</li></ul>
<b>Research Paper Classification Optimization Using Networks and NLP</b>
<ul style="list-style-type: none"><li>• Implemented k-means and SVD to simplify categorization of 2.4+ million arXiv research papers</li><li>• Optimized KNN, random forest, and fastText to develop a high accuracy paper classification model</li><li>• Integrated ANNOY (Approximate Nearest Neighbors) model to retrieve similar papers within 800ms</li></ul>

## SKILLS

<b>Programming:</b>	Python, R, SQL (MySQL, SQLite), JavaScript, HTML/CSS, L <sup>A</sup> T <sub>E</sub> X
<b>Development &amp; Analytics:</b>	Git, AWS (S3, EC2, RDS, Lambda), Databricks, SAS, Excel, Google Analytics
<b>Data Visualization:</b>	Tableau, QGIS, Leaflet, D3.js, Chart.js
<b>Languages:</b>	English and Korean (bilingual fluency)