

# JUNWON CHOI

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

## EDUCATION

### University of California, Davis

Master of Science, Statistics (Data Science Track)

Davis, CA

December 2025

- **Relevant Coursework:** Big Data and High Performance Statistical Computing, Statistical Methods for Research

### University of California, Los Angeles (UCLA)

Bachelor of Science, Data Theory

Los Angeles, CA

June 2024

- **GPA:** 3.7 / 4.0
- **Relevant Coursework:** Data Analysis and Regression, Machine Learning, Data Ethics, Statistical Models in Finance

## EXPERIENCE

### Research Intern

March 2023 – September 2024

Sensing and Robotics for Infrastructure (SRI) Lab at UCLA

Los Angeles, CA

- Led multiple projects to add key features to a robust road assessment web application funded by the City of Los Angeles
- Integrated data processing scripts to AWS S3, EC2 pipeline to automate restructuring of 10+ GB of geospatial data
- Evaluated equity and accessibility of road network using census data and network analysis on 88+ million node pairings
- Customized computer vision models to detect parked vehicles from 150+ GB of onboard camera footage

### Data Team Director / Data Editor

May 2023 – June 2024

The Daily Bruin

Los Angeles, CA

- Directed a team of 20 journalists to publish 50+ articles and data visualization inserts across multiple newsroom sections
- Expanded production capacity threefold by launching two internal production teams and training 15+ new reporters
- Reported GA4 engagement data for all 17 internal teams and advised content strategies to increase overall readership

### Data Journalist

October 2022 – May 2023

The Daily Bruin

Los Angeles, CA

- 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

### Research Team Lead

October 2022 – March 2024

DataRes at UCLA

Los Angeles, CA

- Headed a team of 5+ students to develop machine learning projects based on novel geometric deep learning research
- Increased team's data distribution efficiency by designing custom SQL server to host and filter 50+ GB of raw data

### Sergeant (Communications Squad Leader)

October 2020 – April 2022

Republic of Korea Marine Corps

Gimpo, South Korea

- 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

## RESEARCH/PROJECTS

### Fingerhut (Online Retailer) Strategic Consulting | Capstone Project (Data Theory, B.S.)

- Analyzed 1.6+ million customers' web sessions to extract key features driving account activation and purchases
- Developed a custom stochastic model to predict and visualize customer behavior based on advertisement exposure
- Presented insights and recommended strategic changes to increase customer retention to Fingerhut executives

### Synthetic Survey Data Generation and Evaluation | Research Paper (Submitted to ACM KDD 2025: ADS Track)

- Built end-to-end pipeline to generate synthetic data using U.S. voter files with parametric, GAN, LLM based models
- Assessed synthetic data quality using 4 general utility, 1 target-specific utility, and 4 privacy metrics

### Research Paper Classification Optimization Using NLP and Instance-Based Learning | Course Project

- Developed comprehensive NLP pipeline to process, categorize, and retrieve related research papers from arXiv database
- Compared clustering models at varying complexity against fastText and ANNOY for optimal performance

## SKILLS

**Programming:** Python, R, SQL (MySQL, SQLite), JavaScript, HTML/CSS,  $\text{\LaTeX}$

**Data Technologies:** Git, AWS (S3, EC2, RDS, Lambda), Tableau, Excel, Google Analytics, Hadoop, Databricks

**Machine Learning:** PyTorch, TensorFlow, scikit-learn, NetworkX, igraph, tidymodels

**Languages:** English and Korean (bilingual fluency)