# Jaden Vo (425) 404-1033 | jaden.p.vo@gmail.com | linkedin.com/in/jaden-voo | https://github.com/javomode

# Education

**University of California, San Diego** — B.S. Data Science, Minor in Cognitive Science (ML Specialization)
*Expected June 2027 | GPA: 3.43*

**Relevant Coursework:**
- Data Structures & Algorithms (Java: search/sort, recursion, trees, complexity analysis)
- Software Development Practices (version control, modular programming, testing)
- Foundations of Machine Learning (loss minimization, model optimization)
- Practical Data Science (web scraping, regex, feature engineering, modeling)
- Introduction to Data Management (SQL schema modeling, query optimization)

# Technical Skills

**Languages & Frameworks:** Python (Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn, Plotly, Flask), Java, SQL, Git, REGEX, HTML/CSS (basic)

**Data Engineering:** ETL pipeline design, data cleaning & transformation, schema modeling (ERD, relational databases), query optimization, feature engineering, normalization

**Software Development:** OOP (Java, Python), debugging, testing, version control workflows

# Project Experience

**Recipe Ratings Predictor — UCSD (DSC 80 Final Project),** *May – Jun 2025*

- Designed and implemented **Python-based data pipeline and ML models** on 230K+ Food.com recipes.
- Developed **linear regression and Random Forest models**; improved RMSE from 0.7139 → 0.7075.
- Engineered features (complexity, cooking time categories) and conducted **bias/fairness testing**.
- Applied **modular, reusable code design** for data preprocessing and model evaluation.
- *Tools: Python, Pandas, Scikit-learn, Seaborn, Plotly*

**Trail Difficulty Classification — UCSD (COGS 108 Final Project),** *May – Jun 2025*

- Built **end-to-end ML pipeline** to classify hiking trail difficulty from 3,000+ AllTrails records.
- Applied **feature engineering (gradient, facility count)** and tuned models using GridSearchCV.
- Achieved **81% accuracy with Random Forest**, outperforming Logistic Regression baseline.
- *Tools: Python, Pandas, Scikit-learn, Seaborn*

**Mini ETL Project, Aug 2025 – Present**

- Designed **ETL pipeline** to process raw CSV datasets into a relational schema.
- Implemented transformation logic (handling missing values, normalization, referential integrity).
- Automated loading into SQL database for downstream querying and analysis.
- Applied modular pipeline design and error handling for scalability.
- *Tools: Python (Pandas, SQLAlchemy), SQL*

# Work Experience

**San Diego County Taxpayers Association** — *Data Science Intern* (Database Management & Data Engineering), *Oct 2024 – Jan 2025 | San Diego, CA*

- Compiled datasets from archives and APIs into standardized CSV files for analysis.
- Contributed to **early-stage database planning and automation tools** (scrapers).
- Communicated database scalability/feasibility to leadership, influencing technical planning.

**QWENCH Juice Bar & DRNK Coffee + Tea** — *Front of House Associate*
*Nov 2022 – Jun 2023 | Northridge, CA*

- Improved workflow efficiency by **suggesting and testing process changes**.
- Trained and mentored team members, ensuring compliance with operational standards.

**DC Donut Cafe** — *Front of House Associate*
*Aug 2021 – Sep 2022 | Granada Hills, CA*

- Created a **digital inventory checklist system** to streamline supply tracking.
- Trained staff and supervised closing duties to maintain operational consistency.

---

# Campus Involvement

**UCSD Dragon Boat Team**

- Contributed to **2024 College Cup Championship victory** in San Francisco, CA.
- Strengthened teamwork, communication, and discipline through competitive training.