Weijie (Jack) Zhang

wzhang3912@gmail.com | 747-283-5852 | linkedin.com/in/wzhang3912 | U.S. Permanent Resident

Education

University of California San Diego   Expected Graduation: April 2026
Bachelor of Science in Data Science   GPA: 4.0/4.0

- Coursework: Database Management, Scalable Data System, Data Visualization, Machine Learning, Deep Learning, Data Structure & Algorithm, Probability & Statistics, Optimization, Linear Algebra

Experience

Data Infrastructure Engineer Intern   June 2025 – Present
Boundary Remote Subsurface Solutions   Los Angeles, CA

- Building **ETL data pipelines** in **Python** and **SQL** integrated with **AWS** Lambdas, API Gateway, and S3 to automate **geospatial data** ingestion into a serverless data mesh for ML applications and analytical dashboards
- Engineering and deploying an automated **dataset upload system**, reducing engineering involvement by 70% and accelerating integration of new geospatial data sources for internal teams and external users
- Integrating **data validation** and quality checks into existing data pipelines to ensure reliable ingestion and maintain data governance standards in production

Software Engineer Intern   Mar 2025 – Sep 2025
Green Optic Films   Philadelphia, PA

- Developed a surgical video analysis software application in **Python** (PyQt) to support downstream analytics
- **Collaborated** with cross-functional teams to redesign UI/UX and implement video player and data annotation system for surgical instruments tracking, with detailed technical and design documentation
- Implemented a **CI/CD pipeline** with GitHub Actions to automate development, testing, and packaging for production software distribution, reducing debugging time by **25%** and accelerating software release by **2x**
- **Automated** the extraction and transformation of annotation data into structured formats and built **interactive dashboards** in Tableau to visualize instrument usage trends, supporting reporting and decision-making

Instructional Assistant   Sep 2024 – Present
Halıcıoğlu Data Science Institute at UC San Diego   La Jolla, CA

- Hosting weekly office hours, providing individual and group support to **100+** students for understanding **data science course concepts** (Statistics, Data Structure and Algorithm, Practice of Data Science)

Research & Projects

Research Project: Meta Learning as Planning | *Python (PyTorch, OpenAI's Gym), Docker, Linux Shell*

- **Engineering** software infrastructure with a research team of 5, developing algorithms upon the Decision Transformer models for solving planning problems in Reinforcement Learning
- Building training and inference pipelines in **Python** and **PyTorch**, containerizing environments with **Docker**, and automating workflow via **Linux shell scripting** on computer clusters, reduced experiment setup time by 50%

AI-Driven SQL Database Tool | *Python, PostgreSQL, SQL, Large Language Model, Prompt Engineering*

- **Simplified database access** by building an AI agentic pipeline with a chatbot interface to automatically generate SQL queries for PostgreSQL databases, reducing query-writing **time by 75% per task**
- Utilized prompt engineering techniques and integrated **Large Language Models (LLM) API** with real-time streaming to ensure smooth and robust user experience

Technical Skills

Programming: Python (Pandas, Numpy, Scikit-learn, PyTorch, Dask, Plotly, BeautifulSoup, Selenium, PyQt, PyTest), SQL, Spark, Java, JavaScript, D3.js, HTML/CSS
Technologies: Git, Docker, AWS, Linux Shell, PostgreSQL, Tableau
Concepts: Agile, CI/CD, Prompt Engineering, Infrastructure-as-code, REST APIs, ETL Pipeline