Walter Wong

1024 Geddy Way, Fremont, CA, United States
(510) 304-1861 | wcw003@ucsd.edu | waltercywong.github.io/portfolio/

Education

University of California, San Diego *September 2025 – December 2026*
M.S. in Computer Science

University of California, San Diego *September 2022 – June 2025*
B.S. in Data Science, Minor in Mathematics

Relevant Coursework:
Algorithm Design and Analysis, Parallel Computation, Machine Learning, Recommender Systems, Systems for Scalable Analytics, Deep Learning, Data Management, Data Visualization, Data Science Optimization

Work Experience

Instructional Assistant - University of California, San Diego *September 2023 – March 2025*
Tutored **1000+** students: answered questions during office hours, designed programming assignments, and managed class infrastructure for the courses "Programming and Data Structures" and "Data Structures and Algorithms", which utilize **Java** and **Python** to teach data structures and OOP. Received **100% positive** reviews from students.

Data Science Intern - Resilinc *June 2023 - September 2023*
Designed and implemented a RDBMS to knowledge graph data pipeline for **over 1.5 million** supply chain relationships. Utilized **Apache Spark** DataFrames and Neo4j's Spark Connector to build **knowledge graphs** from schema designs with **12 unique** node and edge types. Wrote queries in **CypherQL** that serve clients worldwide and match SQL queries.

Personal Projects

Document Question Answering Agent *July 2025 – Present*
LangChain, LlamaIndex, Hugging Face Embeddings, ChromaDB, pdfminer, FastAPI, Streamlit
Built a **RAG** system that lets users query PDFs through a conversational interface with source-grounded answers. Deployed with **FastAPI** and **Streamlit** while tuning retrieval and evaluation strategies to improve accuracy.

Spatial and Temporal Modeling for Climate Emulation *April 2025 - June 2025*
PyTorch, Deep Learning Architecture, Convolutional Neural Networks, Transformers, LSTM
Performed data processing, analysis, and visualization on spatial and temporal climate data with over **42.3 million** SSP data points. Developed deep learning architectures that leveraged **CNN**s and **transformers** with up to **38.7 million** parameters that reduced **RMSE** by **34%** from the baseline model.

Congestion Prediction in Chip Design with Walks and Partitioning *September 2024 - March 2025*
PyTorch, Deep Learning Architecture, Graph ML, Analog Chip Design, Chip Congestion
Partnered with Qualcomm industry partners to perform data analysis on **12** unique, congested netlists with **800,000** to **2.6 million** nodes and nets per design. Built on top of **DE-HNN** graph ML model to improve prediction accuracy by **6.7%**. Additionally, built an XGBoost classifier for congested nodes with **92.6%** accuracy on only pre-PnR features.

Visualizing NBA Scoring Trends *February 2024 - April 2024*
Svelte, D3.js, Javascript, HTML/CSS, Pandas, BeautifulSoup, Front-End Framework, Web Design
Created a website with **7** unique interactive visualizations on NBA scoring trends using **Svelte** and **D3.js**. Applied data scraping and cleaning methods on **72 years** of NBA data on **4** different scoring metrics.

Predicting Flight Prices *January 2024 - April 2024*
Scikit-learn, Tensorflow, XGBoost, Pandas, Matplotlib, Classification
Performed airline flight data analysis and visualization to create predictive linear regression, random forest, **XGBoost**, and deep learning models in a robust pipeline with data preprocessing, model training, and performance assessment.

Technical Skills

Languages: Python, Java, SQL, CypherQL, JavaScript, HTML/CSS
Frameworks: Spark, Dask, TensorFlow, JUnit, PyTorch
Developer Tools: Git, VS Code, IntelliJ, Cursor, Tableau, Jupyter
Libraries: Pandas, NumPy, Plotly, Matplotlib, Scikit-learn, D3.js, Scipy, Keras, FastAPI, Streamlit, pdfminer
Cloud Platforms: AWS EC2, Google Cloud (GCP), Microsoft Azure
LLM Tools: LangChain, LlamaIndex, Hugging Face Embedding/Transformer, Ragas, Gemini CLI