

HAO LOU

Los Angeles, CA 90015 | jacoblou0924@gmail.com | (561) 765-0543 | [Linkedin Profile](#)

Education

University of Southern California Agu 2024 – May 2027
Major: Applied and Computational Mathematics(BS) **Minor:** Computer Science(BS) **GPA:** 3.85/4.0

- **Coursework:** Numerical Methods, Probability Theory, Discrete Methods, Calculus III, Data structure, Design theory, math for Machine Learning, **Computer programming:** C++, C, Java, Xcode, R, HTML, matlab, pytorch

Honors

- Platinum Division Qualifier (highest level) in USA Computing Olympic(**USACO**) 2-3% qualification rate
- TOP 30% out of 1805 in **Kaggle research competition:** AMP®-Parkinson's Disease Progression Prediction
- States finalist (top 6 out of 56 teams) in First Technology (**Robotics**) Competition Florida Championship

Research and Work Experience

Large-Scale Occupational Coding & Fuzzy Matching with Prof. Zhuo Chen Sep 2025 – Present
Research Assistant Tsinghua University, China

- Built and iteratively refined a large scale occupational coding pipeline in Python using pandas, jieba, rapidfuzz and Sentence Transformer embeddings to map over 15 million Chinese job titles to 1,636 standardized categories, which improved precision and recall by 18 percent and now serves as the foundation for ongoing clustering, active learning and semi supervised modeling work that will culminate in a co authored manuscript for submission to a peer reviewed journal.

LLM-based Cannabis Use Disorder Detection Research with Dr. Shinyi Wu Dec 2025 – Present
Research Assistant USC, Los Angeles

- Supported an early stage Cannabis Use Disorder auto detection study by labeling several thousand Reddit posts with DSM based symptom and functional impairment guidelines and writing Python scripts for text cleaning, inter rater reliability checks and exploratory model baselines, which will accelerate LLM fine tuning and evaluation.

Huntsman Corporation May 2025 – Aug 2025
product development intern Shanghai, China

- Optimized an industrial methanol synthesis process by building regression and neural network models with multi objective optimization and feature importance analysis to identify an H2 to CO ratio slightly above the theoretical 2 to 1, designing classification models to map inert gas 15 to 25 percent and CO2 2 to 8 percent operating regimes and applying Bayesian and genetic algorithms to explore temperature pressure trade offs, which reduced predicted side reactions, extended catalyst lifetime by an estimated 15 percent and established a foundation for future reinforcement learning based adaptive control.

Kaggle AMP®-Parkinson's Disease Progression Prediction competition Feb 16, 2023 – May 18, 2023
Individual Researcher Online Research competition

- Modeled longitudinal protein and peptide abundance to forecast Parkinson's disease progression via MDS-UPDRS scores using time-series feature engineering on 10,000+ samples, normalization, dimensionality reduction and cross-validated ensemble and neural network models, which achieved a top 30 percent ranking among 1,805 Kaggle teams.

Volunteer and Activities

VOILA - Startup CRM SaaS Sep 2025 - present
Machine Learning Lead Delray Beach, FL

- Increased sync search efficiency by over 80 percent and boosted relevant track hit rate in the top ten results by 30 percent by architecting Viola's text-to-music retrieval pipeline using Python, PyTorch, CLAP embeddings, ChromaDB, emotion-aware re-ranking and evaluation via Precision@k, which powered our investor-ready MVP.

USC Code The Change Nov 2024 - Present
• Teamed Voices Beyond Assault project, implementing a **secure, anonymous** online forum used by **100+** survivors to share experiences and seek support **AND** After the major fire in LA, Build **an interactive platform** to make Blue Sky LA's climate projects visible, searchable, and engaging for the public. Built with **NextJS, CloudFlare, and Mapbox**

Other Projects – Automated Chemistry Lab Workflows

- Automated chemistry lab PDF content extraction, monthly report summarization, MSDS expiration tracking with proactive alerts and procurement forecasting by building integrated workflows in Python (pandas, PyPDF2), VBA macros and scheduled batch jobs, which improved reporting accuracy and reduced manual effort by roughly 40 percent so staff could focus on core lab operations.