CHENYI WENG

F-1 Visa with OPT/STEM OPT Eligibility (Dec 2025 – Dec 2028) Los Angeles, California

✉ wengchen@usc.edu 🏠 mona100421.github.io/chenyi/ 🔗 linkedin.com/in/wengchen/

EDUCATION

University of Southern California   Los Angeles, California
Master of Science, Spatial Data Science (STEM), GPA: 3.71/4.0   Expected Dec 2025
Relevant Coursework: Machine Learning (DSCI552), Data Science at Scale (DSCI550)

National Taipei University of Business   Los Angeles, California
Bachelor of Business of Administration, International Business   Jan 2022
Minor in Information Management
Relevant Coursework: Object-Oriented Analysis, Data Structures, Operation System, Computer Security

PROJECTS

USC Viterbi School of Engineering
DSCI552 Machine Learning   Los Angeles, California
Supervised & Classical ML   May 2025 to Aug 2025
- Implemented baseline models (Logistic Regression, Decision Trees, Random Forests, SVM, kNN) for structured datasets.
- Conducted data preprocessing, normalization, and cross-validation to improve model generalization.

Ensemble Learning& Active Learning Frameworks
- Applied ClassifierChain for multi-label classification and evaluated with classification reports & confusion matrices.
- Developed Active Learning SVM approach: selected informative samples closest to the decision boundary, improving learning efficiency.

Deep Learning Pipeline for Image and Audio Classification
- Designed and trained CNNs for image classification tasks; compared against classical ML baselines.
- Final Project: built full ML pipeline for audio dataset (data preprocessing → feature extraction → model selection → hyperparameter tuning).
- Achieved >95% accuracy, demonstrated robustness through ROC AUC and F1 metrics, and documented methodology in reproducible GitHub repo.

DSCI550 Data Science at Scale   Los Angeles, California
Scalable Data Extraction Pipeline   Jan 2025 to May 2025
- Implemented a Python + Spark pipeline to extract, clean, and transform large-scale unstructured data.
- Used regex and batch processing to improve parsing efficiency and achieved >90% clean data accuracy.

Interactive Data Visualization Dashboard
- Built interactive dashboards with Plotly/Dash to visualize million-scale datasets.
- Added filtering and drill-down functionality to support data exploration and decision-making.

TECHNICAL SKILLS

Programming & Framework: Python (Flask, FastAPI, NumPy, Pandas, scikit-learn, TensorFlow, Keras), JavaScript (React, Node.js, Express), TypeScript, SQL/NoSQL (PostgreSQL, MySQL, MongoDB), Go, C, Java, HTML/CSS
Systems & Tools: Git, GitHub, Bash, CI/CD (Docker, GitHub Actions), AWS, Testing & Debugging (Pytest)
Concepts: Full-Stack & Web Application Development, Cloud Computing, Scalable Systems & Performance Optimization, Security & Compliance (KYC, Fraud Prevention), Data Analytics & Machine Learning, API Design
Languages: Mandarin (native), English (fluent)