

Yuqiong “Jade” Liang

✉ jade.y.liang@gmail.com
in linkedin.com/in/jade-yuqiong-liang
github.com/jade-y-liang
🌐 jadeyliang.com

EDUCATION

University of California, Los Angeles

Dec 2025

B.S. Mathematics and B.S. Statistics & Data Science, Specialization in Computing

- Department GPA: 3.9
- Related Coursework: Python with Applications, Probability and Statistics, Linear Models, Data Analysis and Regression, Design and Analysis of Experiment, Data Structures and Algorithms, Linear Algebra.

EXPERIENCE

Kurma AI

Remote

Data Analyst Intern

Feb 2025 - Present

- Build automated pipelines to extract raw text data and embedded images (e.g. tables, plots, figures) from research paper PDFs to train RAG and LLM models.
- Clean extract text by removing special characters, citation patterns, and formatting inconsistencies, and implement code to exclude irrelevant images (e.g., logos, blank images).
- Research and evaluate data extraction and cleaning methods to identify the most effective techniques for producing model-ready data.

Wing Cheung Co.

San Francisco, CA

Cashier

Jun 2015 – Aug 2018

- Provided friendly, accurate service to hundreds of customers during daily transactions.
- Maintained organized, customer-friendly produce displays throughout each shift.
- Memorized and recalled prices for 35+ items to ensure fast, accurate checkout.

SKILLS

Languages Python, R, SQL, C++, $\LaTeX$.

Technical Git, NumPy, pandas, Matplotlib, seaborn, Plotly, ggplot2, reactable, sqlite3, Scikit-learn, Keras, Tensorflow, Dash, Quarto, Scrapy, Power BI, Microsoft Excel.

PROJECTS

No-Plan Pantry

Oct 2024 – Dec 2024

- Built a web scraper using *Scrapy* to extract 15.2k recipes from allrecipes.com for model training.
- Transformed raw data using *pandas*, and tokenized recipes to train *GPT-2*.
- Used *SQL* to extract recipes data based on calorie range, ingredients, and time limit.
- Fine-tuned *GPT-2* on curated dataset to generate recipe suggestions tailored to user-inputted ingredients.
- Designed and accomplished backend logic for *Dash* web application.

Image Classification of Cats and Dogs

Nov 2024 – Dec 2024

- Created neural network models using *keras.sequential* to distinguish between images of cats and dogs with 96.73% accuracy, achieving an accuracy increase of 46.53% over baseline model.
- Executed strategies such as data augmentation and data preprocessing to enrich training datasets, and applied transfer learning using a pre-trained *keras.sequential model* to revamp learning efficiency.

Fake News Classification

Nov 2024 – Dec 2024

- Standardized text data and optimized model training by batching dataset to process in chunks rather than individual rows, which enhances computational efficiency.
- Designed and trained binary classification models using *Keras* to classify news articles as real or fake.
- Fine tuned model parameters to attain 97.54% accuracy, producing a 45.2% accuracy increase over baseline model.