Yuqiong “Jade” Liang

[redacted] | jade.y.liang@gmail.com | jadeyliang.com | github.com/jade-y-liang | linkedin.com/in/jade-yuqiong-liang

EDUCATION

University of California, Los Angeles (UCLA) *Expected Graduation: Dec 2025*
Bachelor of Science in Mathematics, Statistics and Data Science (Double Major) *Department GPA: 3.9*
Specialization in Computing

- **Relevant Coursework**: Python with Applications I and II, Probability and Statistics, Linear Models, Data Analysis and Regression, Design and Analysis of Experiment, Data Structures and Algorithms, Linear Algebra.

EXPERIENCE

Data Analyst Intern | *Kurma AI*, Remote *Feb 2025 - Present*

- Assisted in the preparation and cleaning of PDF documents to train Large Language Models (LLMs) and Retrieval-Augmented Generation (RAG) models.
- Explored innovative text extraction methods to enhance data formatting and quality for machine learning applications.

PROJECTS

No-Plan Pantry (Recipe Recommendation System) | *pandas*, *Scrapy, SQL, Dash, GPT-2* *Oct 2024 – Dec 2024*
Link: https://github.com/jade-y-liang/no-plan-pantry

- Constructed a web scraper using *Scrapy* to extract 15.2k recipes from allrecipes.com for analysis and model training.
- Cleaned and transformed raw data using *pandas,* and tokenized recipes to train GPT-2.
- Used *SQL* to properly 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, and devised callback functions to process user-inputted ingredients dynamically.

Image Classification of Cats and Dogs | *Keras, Tensorflow, pandas* *Nov 2024 – Dec 2024*
Link: https://jade-y-liang.github.io/pic16b-blog/posts/hw5-image-classification

- 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 | *Keras, Tensorflow, pandas* *Nov 2024 – Dec 2024*
Link: https://jade-y-liang.github.io/pic16b-blog/posts/hw6-real-vs-fake-news

- Standardized text data by converting to lowercase and removing stop words from titles; 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 *tensorflow.keras.losses.Binaryrossentropy* to accurately classify news articles as real or fake, and integrated both article titles and text content as model inputs.
- Fine tuned model parameters to attain 97.54% accuracy, producing a 45.2% accuracy increase over baseline model.

Academic Performance Analysis and Predictor | *R* *Jul 2024 – Aug 2024*
Link: https://github.com/jade-y-liang/academic-performance-analysis

- Conducted comprehensive data analysis on a dataset of 10,000 student responses.
- Identified key factors correlating with academic performance index through rigorous statistical analyzing, using *R* to handle data cleaning, manipulation, visualization and analysis.
- Constructed and refined a predictive model using linear regression techniques to forecast student performance indices.
- Predicted student Performance Index by fine-tuning a linear regression model (checking base model, model after power transformation, and model after variable selection), which reduces overfitting from baseline model.

SKILLS

Languages: Python, R, SQL, C++, LaTeX.
Other Software/Tools: Git, NumPy, pandas, Matplotlib, seaborn, Plotly, ggplot2, reactable, sqlite3, Scikit-learn, Keras, Tensorflow, JAX, Dash, Quarto, Scrapy, HTML.