MEIYI YE

Meiyiye123@gmail.com · https://www.linkedin.com/in/meiyiye/ · https://meiyi-ye.netlify.app/
626-757-9062 · https://github.com/Meiyi-Ye

EDUATION

University of California, Los Angeles (UCLA) — Los Angeles, CA
Bachelor of Science in Statistics and Data Science — June 2024

- Relevant Coursework: Probability / Mathematical Statistics / Linear Models / Data Analysis and Regression / Experimental Design / Computational Statistics with R / Statistical Programming with R / Python / Computation and Optimization / Statistical Models and Data Mining / Monte Carlo Methods / Statistical Consulting

SKILLS

Technical: R: tidyverse (dplyr / ggplot / readr), Python: scikit-learn / nltk / keras / TensorFlow / PyTorch / pandas / numpy / scipy / statsmodels / matplotlib / seaborn, SQL, Tableau, Microsoft Excel, Statistical Modeling, Machine Learning, Time Series, Natural Language Processing, HTML, CSS, GIS, Github

Languages: English, Mandarin, Cantonese

EXPERIENCE

BYD — Lancaster, CA
After-sale Service Support Engineer — Sep 2024 - Present

- Working under the BESS (Battery Energy Storage System) Department.

Bluebonnet Data — Los Angeles, CA
Data Fellow — Aug 2023 - Nov 2023

- Developed **interactive maps** and **visualizations** in **R**, **Python**, and **Tableau** to track donor metrics, improving stakeholder decision-making by 25%.
- Analyzed 150,000+ donation records (2014-2022), uncovering trends that improved donor engagement by 15%.

UCLA — Los Angeles, CA
Student Researcher — Jan 2023 - Mar 2023

- Developed and implemented **CNN** models in **Python** for image recognition, predicting human performance with 85% accuracy in online experiments.

Antelope Valley College — Lancaster, CA
Student Instructor — Jun 2021 - May 2022

- Tutored over 100 college students in **C**, **C++**, and Statistics resulting in over 78% of tutees receiving an “A” in the subject
- Held group and one-on-one sessions as necessary, and received over 97% positive feedback from students on competence and approachability

PROJECTS

IMDB Reviews Sentiment Prediction

- Utilized **Logistic Regression**, **KNN**, **LDA**, **QDA**, and **Random Forest** models in **R** to predict binary sentiment based on word frequencies from 50,000 movie reviews.
- Implemented 10-fold LDA model, achieving a testing accuracy of 73.52%, the highest among all models.

Yelp Review Rating Prediction

- Built models using **CNN**, **LSTM**, and **BERT** in **Python** to predict true star ratings from 7 million+ reviews.
- Achieved an MSE of 0.4, with predicted ratings typically 0.63 stars away from the actual rating.