PINYI LI

+1(805)-679-1295 | pli216190@gmail.com | Los Angeles, CA, 90024

EDUCATION BACKGROUND

University of California, Los Angeles June 2023 - June 2025 (Expected)
B.S. in Chemistry (Major) & Statistics (Major) | **GPA: 3.96/4.00** Los Angeles, CA
University of California, Santa Barbara Sept 2021 - June 2023
B.S. Chemistry | **GPA: 3.98/4.00** Santa Barbara, CA

Relevant Coursework: Statistical Modeling, Data Mining, Computational Statistics, Monte Carlo Methods, Design and Analysis of Experiment, Linear Models, Data Analysis and Regression, Probability and Statistics, R Programming, C++, Python, Real Analysis…

RESEARCH EXPERIENCE

University of California, Los Angeles Jun 2024 – Oct 2024
Undergraduate Researcher with Professor Xiaowu Dai Los Angeles, CA

- Conducted research inspired by Doyle Group's study *"Identifying General Reaction Conditions by Bandit Optimization,"* focusing on optimizing chemical reactions condition using machine learning techniques.
- Tested a broader range of parameters and algorithm variations (e.g., epsilon-greedy, UCB, Thompson Sampling, Bayesian UCB) under simulated scenarios, then applied findings to identify optimal reaction conditions.
- Enhanced the accuracy in identifying optimal ligands for C-H arylation to 88% with a new tested variation. Visualized results to compare algorithm performance.
- Performed in-depth regret analysis on algorithm outcomes using Python and R, and visualized results.

University of California, Santa Barbara Sept 2022 - June 2023
Research Assistant in Professor Arnab Mukherjee's Lab Santa Barbara, CA

- Prepared essential growth media and reagents including RPMI, agar gel, and DMEM for cellular experiments, ensuring precise formulation and quality control.
- Performed tissue culture operations such as cell maintenance, splitting, and cryopreservation; executed advanced molecular cloning techniques including primer design, PCR, gel electrophoresis, gel extraction, Gibson assembly, bacterial transformation, and plasmid preparation (mini/midi prep).

PROJECT EXPERIENCE

Statistical Models and Data Mining Class Group Project July 2024 – Aug 2024
Regression Project: Amazon Order Cost Prediction

- Developed a predictive cost estimation model in R for Amazon orders, conducting exploratory data analysis and visualizing key trends to inform feature selection and model design.
- Preprocessed and cleaned data to ensure model accuracy; implemented multiple machine learning algorithms, including linear and lasso regression, random forest, SVM, XGBoost, and neural networks.
- Optimized model performance through hyperparameter tuning and applied model stacking techniques; evaluated models using RMSE and presented comparative results through visualizations.

Classification Project: 2020 United States Presidential Election Winner Prediction

- Designed and developed a predictive model to forecast the winner of the 2020 US Presidential Election, utilizing exploratory data analysis to uncover and visualize key trends for improved insights.
- Preprocessed datasets and implemented a range of machine learning algorithms, including logistic regression, random forest, SVM, and KNN, to create a robust predictive framework.
- Tuned hyperparameters and applied model stacking techniques to enhance predictive accuracy and visualized comparative results to effectively assess model performance.

Design and Analysis of Experiment Class Project Apr 2024 – June 2024
Impacts of Dark Chocolate with Varied Cocoa Content and Dosage on Urine Dopamine Levels

- Led a randomized complete block design study to assess the impact of dark chocolate's cocoa content (40%, 70%, 85%) and dosage (50g, 100g, 150g) on urine dopamine levels in a simulated island population, using two-way ANOVA for data analysis.
- Managed all aspects of experimental design, including multi-stage sampling, data collection, and model validation (residual plot, interaction plot), while identifying opportunities for improving the procedure and controlling for confounding factors.

Data Analysis and Regression Class Project Jan 2024 - Mar 2024
Investigation of Alcohol Consumption's Impact on Liver Disorder

- Developed and refined a linear regression model to explore the relationship between blood test indicators of liver disorders and alcohol consumption, using backward stepwise regression for variable selection.
- Validated model accuracy by assessing multicollinearity with variance inflation factor (VIF) and residual plot analysis.

PROFESSIONAL EXPERIENCE

Data Analyst Intern, **Deloitte AI Institute,** Shanghai, China Aug 2023 – Sept 2023

- Evaluated AI model performance across general knowledge, logical reasoning, and translation tasks, documenting accuracy metrics, identifying improvement areas, and compiling pricing data to support commercial deployment strategies.
- Tested and analyzed nearly 100 AI plugins for office applications, assessing functionality, requirement alignment, and user experience. Provided detailed feedback on key features, input-output formats, and potential issues for real-world usage.

EXTRACURRICULAR & LEADERSHIP

Teaching Assistant for Chemistry Lab Course, **UCSB Department of Chemistry & Biochemistry** Sept 2022 - Dec 2022

- Partnered with teaching assistants to explain key concepts to students, enhancing their understanding and engagement.
- Guided students through experimental procedures, lab report preparation, and post-lab assignments.

Teaching Volunteer in Middle School, **UCSB SciTrek** Sept 2021 - Dec 2021

- Conducted hands-on experiments to educate students on concepts of the conservation of mass and closed/open systems
- Guided students to design and conduct experiments to validate learned concepts

ADDITIONAL INFORMATION

- **Technical:** R, C++, Python, MATLAB, PyMOL, Microsoft Office (Word/Excel/PowerPoint)
- **Language:** Chinese (native), English (proficient)