KLOE YAP

📞 (510) 316-4080 ◇ Fremont, CA
✉ kloeyap@berkeley.edu 🔗-kloeyap

EDUCATION

University of California, Berkeley — **May 2024**

Bachelor of Arts, Data Science, Domain Emphasis: Social Law and Policy — *GPA: 3.93*

Relevant Coursework: Foundations of Data Science, Data Structures and Algorithms, Principles and Techniques of Data Science, Data, Prediction, and Law, Data Engineering, Probability for Data Science, Data, Inference, & Decisions, Natural Language Processing, Applied Econometrics

SKILLS

Languages: Python, R, SQL, MongoDB, Java, C++, MATLAB
Libraries & Visualization Tools: Pandas, Seaborn, Matplotlib, Plotly, Dash, Tableau, Power BI
Data Management: AWS Cloud Services, DBeaver, Snowflake
Machine Learning & Statistical Analysis: Scikit-learn, PyTorch, TensorFlow, NLP, Bayesian & Frequentist Statistics

EXPERIENCE

Data Science Intern — **September 2023 – December 2023**

Allbirds - Data Science Discovery Program — *Remote/San Francisco, CA*

- Initiated a project harnessing the capabilities of the NVIDIA ecosystem, integrating technologies like NVTabular for extensive big data processing which drove 10-20% surge in revenue associated with email interactions
- Executed the deployment and optimization of a Deep Learning Recommendation Model (DLRM) within a recommendation framework achieving 0.97 recall
- Led the revamp and standardization of a recommendation framework, streamlining its process for filtering recommendations and facilitating efficient deployment with an improvement of more than 20% in its accuracy compared to its current model

Data Analyst/Consultant — **September 2023 – May 2024**

Berkeley Core Consulting Group — *Berkeley, CA*

- Developed a systematic week-by-week plan to guide digital record label in increasing awareness of their music among specific demographics, aligning with the overall business objective of contributing to a strategic partnership with a nearby coffee shop
- Revamped a low-profile lo-fi artist discovery by conducting in-depth analysis of streaming data and successfully increased artist awareness and listening activity by 30%
- Employed web scraping techniques on Google Maps and utilized Yelp API to gather 1000 coffee shops in college towns to identify a diverse array of mom-and-pop shops for the record label to reach out to

Data Science Intern — **June 2023 – August 2023**

Aon IP Solutions — *Remote*

- Led back-end development for a dynamic web application facilitating user-defined competitive analysis of companies, sectors, and patent portfolios, utilizing a proprietary scoring system for effective ranking
- Designed and implemented an intuitive web-based tool capable of processing uploaded files or manual inputs to generate numerical scores, aiding in entity positioning within user-defined spaces
- Conducted XML parsing, capturing 95% of relevant information, to analyze patents filed under the medical devices category to enable effective numerical data analysis and pattern recognition

PROJECTS

Making Predictions & Causal Effects from Greenhouse Gases. *Bayesian & Frequentist Methods, Causal Inference*
Applied multi-linear regression, decision tree, and random forest models to analyze socioeconomic factors impacting CO_2 emissions. Utilized causal inference and outcome regression techniques to estimate the causal effect of CO_2 emissions on life expectancy.

Predicting Recidivism. *Python, Pandas, Seaborn, Matplotlib, Scikit-learn*
Employed various techniques to balance data classes, identified optimal features, and applied logistic regression and random forest classifiers to predict defendant flight risk at arraignment hearings. Simulated judge decision-making processes to enhance model accuracy.