

Alexander Bloyer

bloyer.alexander@gmail.com

(408) 427-5855

Gilroy, CA

github.com/alexbloyer

EDUCATION

University of California, Santa Barbara

B.S., Statistics and Data Science

Sept 2022 – June 2025

3.92 GPA

EXPERIENCE

Data Intern

Jan 2025 – June 2025

NAI Northern California

- Developed web-scraping python script to cross-reference obituary data with the company Salesforce database to explore property leads. Leveraged Playwright & Selenium for browser automation.
- Merged +18,000 duplicate contact records in Salesforce database, decreasing render times for agent call lists
- Performed weekly data loads via DemandTools, creating dashboards to convey KPIs and agent performance

Website Chair

May 2024 – June 2025

UCSB, Data Science Collaborative

- Performed weekly website updates with Quarto, HTML, & CSS via Github Pages. Implemented google analytics to track traffic and improve site visitation
- Coordinated staff meetings to plan data science workshops, fundraising events, & community socials
- Hosted and participated in EOY project showcase with NBA subreddit sentiment analysis project.
- Provided statistics/programming advice to students at weekly project showcase workshops & information sessions

PROJECTS

Resume Optimizer (Flask/Python)

- Developed full-stack Flask web application

Sentiment Analysis: NBA Subreddits (Python)

- Gathered reddit user comments across popular NBA subreddits via PRAW, processed text data with custom RegEx function, and analyzed sentiment with NLTK.
- Performed EDA, graphing histograms and word clouds exploring term frequency and sentiment patterns across subs
- Fit a Logistic Regression, Random Forest, and Linear SVM to the vectorized comment data achieving a 75.57% test accuracy score with the Logistic Regression model

Inferential Machine Learning: Steel Plate Fault Analysis (R)

- Utilized UCI Machine Learning Repository for dataset, cleaning feature data and transforming response variables with one-hot encoding
- Fit Bagging and Random Forest models achieving respective test error rates of 18.25% and 21.34%.
- Examined variable importance via accuracy and Gini index to identify significant components of the tree-based methods and discover trends. Found that the conveyor belt length and the plate thickness were the most significant indicators for identifying faults.

Predictive Modeling: Diamond Price Analysis (R)

- Performed EDA, fit multiple linear regression models, performing forward and backward regression to generate prediction models, testing model accuracy with BIC and MSE. Transformed response and feature variables to address non-linearity and heteroscedasticity, improving final model accuracy by 10% with adjusted R-squared 0.9663
- Collaborated with a team using GitHub for version control to manage code and workflows, meeting project deadlines while ensuring group cohesion and effective task delegation

SKILLS

Data Analysis & Visualization, Natural Language Processing, Statistical Modelling, Data-driven problem solving

FRAMEWORKS/TOOLS

Python, R, SQL, HTML/CSS, Git/GitHub, Excel, Visual Studio Code, Jupyter Notebooks, Google Colab, R Studio, NumPy, Pandas, Matplotlib, Seaborn, Altair, Scikit-Learn, Pytorch, Flask, NLTK

RELEVANT COURSEWORK

Python Programming, Statistical Machine Learning, Bayesian Statistics, Time Series Analysis, Base SAS Programming, Regression Analysis, Probability & Statistics Theory, Linear Algebra, Data Science Principles in R & SQL, Design of Experiments, Discrete Stochastic Processes, Single/Multivariate Calculus, Differential Equations, Intro to Mathematical Proofs