Matthew James Lim

matthewlim@berkeley.edu | (925) 719-6872 | linkedin.com/in/matthewjlim25

EDUCATION

University of California, Berkeley *Expected Graduation: May 2025*
B.A. Data Science & Business/Industrial Analytics

- **Relevant Coursework:** Principles & Techniques of Data Science, Probability for Data Science, Data Structures C++ / Java, Object Oriented Programming C++, Multivariable Calculus, Linear Algebra & Differential Equations, Discrete Mathematics, Ethics of Data, Engineering Statistics and Forecasting
- **Activities:** Data Science Society Project Manager, Theta Tau Professional Engineering Fraternity

WORK EXPERIENCE

Snap Inc. **Berkeley, CA**
Data Science Contractor *Sep 2023 – Dec 2023*

- Web scraped user post data from Twitter and Reddit, identifying trends in marginalized communities
- Employed NLP techniques to analyze and identify recurring patterns in the collected data, facilitating the aggregation of trends from multiple sites in a user-friendly and accessible Plotly dashboard

PricewaterhouseCoopers (PwC) **San Francisco, CA**
Consulting Solutions Intern *Jun 2023 – Jul 2023*

- Delivered a client strategy for CareerVillage, ranking top 7% (6/77) of PwC's national consulting interns
- Proficiently analyzed a comprehensive dataset containing 37,000+ rows, employing a synergistic approach that combined Excel for data preprocessing, Python natural language processing techniques (e.g. sentiment analysis and text classification) for content mining, and Microsoft Power BI for creating interactive data visualizations, facilitating data-driven strategies and informed decision-making

PROJECT EXPERIENCE

Music Mate | *Python, Pandas, Seaborn, Scikit Learn, Random Forest* *Feb 2023 – Present*

- Deployed an interactive Flask website that analyzes user Spotify listening habits through song insights
- Built a mood classification Random Forest model, analyzing key song features through a machine learning pipeline including EDA, feature engineering, hyperparameter tuning, and model evaluation

Smart Environmental Monitoring System | *Python, Pandas, Flask, Plotly, Arduino* *Oct 2023 – Nov 2023*

- Implemented various sensors to monitor environmental factors (AQI, radiation, etc.) via Raspberry Pi
- Led backend development for real-time analysis of sensor data in an interactive Plotly dashboard

Predicting Housing Prices | *Python, Regex, Scikit Learn, Regularization* *Oct 2023*

- Conducted housing price prediction analyses utilizing feature engineering, regression, and CV
- Achieved <200k in testing RMSE and ~90% accuracy across 500,000+ Cook County housing records

Plagiarism Detector Using Rabin-Karp Algorithm | *C++, Hash Maps* *Nov 2022*

- Engineered a plagiarism detection software in C++, leveraging the Rabin-Karp algorithm and hash maps for optimizing pattern recognition within text, significantly reducing computational complexity
- Conducted rigorous testing and debugging to ensure robust performance against various text inputs

Wordle Game Implementation | *C++, Object-Oriented Programming* *April 2022*

- Implemented the Wordle game in C++, focusing on object-oriented programming conventions
- Crafted an interactive console-based user interface, allowing for intuitive gameplay interactions
- Integrated a dynamic word selection mechanism by parsing valid words from an external text file

SKILLS & INTERESTS

Skills: Python, Java, SQL, C++, Pandas, Machine Learning, Object-Oriented Programming, Matplotlib, Seaborn, Scikit-learn, NLTK, Tensorflow, Excel, Adobe Creative Cloud, Decking, Data Visualization

Interests: Jazz Saxophone, Espresso Brewing, Weightlifting, DJing, Reading (e.g. Zero to One)