Jeff Hwang

714-262-1742 | jeffhwang9@gmail.com | linkedin.com/in/jeffhwang9 | github.com/JeffyWongo | jeffhwang.dev

EDUCATION

University of California, Riverside   Riverside, CA
Bachelor of Science in Data Science   *December 2024*

EXPERIENCE

Data and Analytics Intern   Jan 2024 – Aug. 2024
Traffic Control Engineering   *Brea, CA*

- Coordinated with the internal engineering team members to design, develop, and maintain dashboards and reports with PowerBI to track key business metrics and support decision-making.
- Collected, cleaned, and organized data for analysis, identifying trends, patterns, and correlations using SQL and Excel
- Conducted thorough reviews of datasets and analyses during project phases, while maintaining and updating databases and logs to ensure data integrity and accessibility.

PROJECTS

Marketing Analysis | *Microsoft SQL Server* | *Python* | *pandas* | *PowerBI*

- Conducted a comprehensive analysis of conversion rates, identifying key drop-off points in the customer journey that improved the conversion rate from **5.0%** in October to **10.2%** in December.
- Utilized sentiment analysis and thematic clustering on customer feedback, uncovering recurring issues in products rated below 3.5 and offering insights to enhance product quality and customer satisfaction.
- Performed an in-depth study of customer engagement metrics, including click-through rates **(15.37%)**, to identify high-performing content types and optimize marketing strategies.

Adult Income Predictor | *Python* | *pandas* | *NumPy* | *scikit-learn* | *tensorflow*

- Developed an income prediction model using the UCI Census dataset, following an end-to-end data science pipeline including data collection, cleaning, preprocessing, and analysis.
- Utilized logistic regression and custom prediction models to forecast income ranges, achieving alignment with underlying data distribution
- Implemented bias mitigation strategies, including resampling, regularization with dropout layers, and fairness analysis using distribution comparisons.
- Evaluated fairness within predictions, identifying limitations due to dataset imbalance and low high-income representation.

Aerosol Concentration Predictor | *Python* | *pandas* | *NumPy* | *Matplotlib* | *scikit-learn* | *seaborn*

- Built a machine learning model to analyze the relationship between weather conditions and atmospheric aerosol concentrations in the Los Angeles, California region
- Conducted KNN regression and used cross-validation to test KNN regression model with **.98** accuracy and **.002** MSE for predicting NOx levels
- Applied data cleaning and exploratory data analysis on a **2000+** observations dataset using tools such as **pandas, seaborn,** and **scikit-learn**

Wine Quality Analysis | *R* | *ggplot2* | *dplyr* | *tidyr* | *ggcorrplot* | *car*

- Developed interaction and second-order models to predict wine quality based on variables such as fixed acidity, volatile acidity, residual sugar, and alcohol content.
- Explored the relationship between chemical composition and sensory quality through laboratory measurements and blind sensory evaluations.

TECHNICAL SKILLS

Languages: Python, SQL, R, C++, JavaScript, HTML, CSS
Tools/Technology: PowerBI, Microsoft SQL Server, Excel, pandas, Numpy, Matplotlib, scikit-learn, tensorflow, seaborn, Git, LaTeX