Wajahat Khan

Bay Area, CA | 510 - 309 - 7734 | wajakhan@berkeley.edu | Github Projects | LinkedIn

Data Scientist

Pressure-resistant data science strategist leaning into curating a more prosperous and sustainable world through technology. Employing a 'roll-up-my-sleeves' work ethic to challenge inefficiencies, apply data manipulations with logic, and collaborate with diverse teams to execute cutting-edge digital projects.

Leadership

- **Awards:** Harvard-Certificate Selection for Award • National Academy of Future Scientists & Technologists Award of Excellence • UC Berkeley Official Paleontologists Assistant Award • Phi Beta Kappa

Technical Aptitude

Coding Languages: Python • Java • R • SQL • Excel

Data Science: Exploratory Data Analysis & Data cleaning • Pandas • Text Wrangling & Regex • Sklearn • Gradient Descent • Feature Engineering • Regularization Random Variables • Cross Validation • Logistic regression • PCA • KMeans clustering • Decision Trees • Linear algebra • Machine Learning • Github • API • Large Data • API

Certificates: Chevron ROP, Data Science & Business Analytics, Generative AI

Industry Assets: Programming Languages • Communication • Modeling & Analytics Technologies

Work Experience

Data Science Internship | The Spark Foundation | Singapore | Dec 2022-May 2023

Curated a progressive environment to engage in a rigorous Graduate Rotational Internship Program to retain real-world data science applications experience.

- Established workflows to execute **9** multifaceted team projects with the Data Mining team and Machine Learning team to transform data into insights and create innovative problem-solving solutions.
- Applied agile frameworks to conduct thorough data analysis of the USA map using Python and SQL, retrieving data insights impacting salary in correlation to religion, state, and city, resulting in a **20%** increase in data accuracy.
    - This data generated analysis to support businesses with the appropriate information for investments and ROI.

Data Entry/Information Technology/Assistant | Lita Pettus-Dotson Law Firm | Alameda, CA | Jun 2018-Aug 2018

Amassed the end-to-end data management transformation from paper to electronic storage eliminating printers and legal disposal services to scale sustainability.

- Implemented portable USBs that drastically reduced cost and time by **50%** and advocated for eco-friendly cost-effective data management.
- Developed and managed data architectures streamlining data integrity and maintaining a **99%** accuracy rate.

Project Snapshot

UC Berkeley Bike Sharing | University of California, Berkeley | Remote | Aug 2023-Dec 2023

Initiated and spearheaded the bike-sharing project managing a team of **7** to predict demand, optimize model variables, exceed milestones, and escalated project efficiency by **20%**.

- Managed the creation of a top-tier model from scratch successfully outperforming master student benchmarks.
- Achieved **99.9%** accuracy by utilizing a random forest model to predict a given demand for bikes for the day.

Kaggle Competition-Titanic | University of California, Berkeley | Berkeley, CA | Aug 2023-Dec 2023

Operated as the Project Manager to lead a team in the Kaggle competition against Master's students applying machine learning algorithms containing random forests and logistic regression to hit a **Top 10%** ranking out of **700 participants**.

- Leveraged Google Collab for a seamless partnership during a two-week project, predicting survival outcomes based on the following factors: gender and age.

Education

- **Master of Science, Data Science & Analytics |Eastern University | St Davids, PA |August 2024 — Present**
- **Bachelor of Science, Data Science| University of California, Berkeley | Berkeley, CA | August 2019 - Dec 2023**