

Ethan Vaz Falcao

Greater Boston

508-808-3510

Ethan.falcao@gmail.com

ethanfalcao.github.io/

linkedin.com/in/ethan-falcao/

EDUCATION

Worcester Polytechnic Institute

Bachelor of Science in Data Science & Mathematical Sciences

05/2024

Minor in Computer Science

Relevant Coursework: Machine Learning, Algorithms, Principles of Real Analysis, Linear Algebra, Combinatorics, Probability Theory, Mathematical Statistics, Artificial Intelligence, Big Data, Business Intelligence

Awards: Pi Mu Epsilon (U.S. Mathematics Honors Society)

Admitted: 04/2023

EXPERIENCE

Schneider Electric

Boston, MA

Data Science Intern

06/2023 - 09/2023

- Constructed a predictive model with Python (PySpark) and random forest regression for quantitative analysis.
- Modeled an impressive 91% accuracy in accurately forecasting UPS Phase 1 sensor battery percentages.
- Successfully uncovered critical insights into battery aging, chemistry, and effective energy conservation over time.

United Solutions

Leominster, MA

Data Analyst Intern

06/2022 - 09/2022

- Designed interactive Power BI dashboards with dynamic DAX formulas, reducing report generation time by 30%.
- Enhanced data visualization techniques facilitated quicker, more informed decision-making based on key trends.
- Implemented daily IQMS Reports in Excel, cutting time and offering insights into labor productivity trends.

Clora

Boston, MA

Business Analytics Intern

06/2021 - 08/2022

- Conducted payroll analysis for 100+ healthcare consultants, ensuring 100% accuracy and contract compliance.
- Developed monthly financial reports, identifying cost-saving opportunities, leading to a 5% budget reduction

PROJECTS

Mining Graph Patterns from Github Repositories | Python, NetworkX, PowerBI, Git

github.com/EthanFalcao/GraphMiningMQP

2023/08 - Present

- Developed ML algorithms for Github code quality assessment: 95% accuracy, analyzing over 900+ repositories.
- Integrated sophisticated features to detect and calculate cyclomatic complexity and Halstead effort
- Designed an algorithm to categorize 1,000+ GitHub repositories as Active, Dormant, or Inactive

Optimizing Crypto Options Models | Deribit API, Crypto Options, Black-Scholes model

github.com/EthanFalcao/Optimizing-Options-Models-Crypto

02/2024

- Developed Bitcoin and Ethereum options models by extracting 12 months of future data from the Deribit API.
- Improved accuracy by 15% through adapting the Black-Scholes model to crypto's unique volatility.

Sentiment Analysis in Financial Markets | ApifyClient, yfinance, NLTK, Scikit-learn

github.com/EthanFalcao/Sentiment-Analysis-in-Financial-Markets

02/2024

- Utilized advanced NLP, including sentiment analysis and topic modeling, to analyze 10,000+ CNBC articles.
- Achieved a notable 85% accuracy rate in predicting stock trends from nuanced market sentiment.

Full Stack - Spotify Recommendation Algorithm | Spotipy, Random Forest, LightGBM

github.com/EthanFalcao/Spotify_Recommendation_Algorithm

12/2023

- Developed a playlist recommendation algorithm with Spotify API data using random forest and cosine similarity.
- Achieving 90% accuracy in personalized playlist recommendations from user-input playlists.

Yelp Dataset Challenge: Detecting Fraud reviews and Business Attributes | NLTK, Pandas, Scikit-learn

ethanfalcao.github.io/Yelp_dataset_Challenge

05/2023

- Accurately predicted business attributes with 86% accuracy, offering deep insights into diverse business profiles.
- Successfully implemented fraud detection, flagging 20% of fake reviews for significantly enhanced data integrity.

LEADERSHIP

Data Science Club Treasurer

Worcester, MA

Worcester Polytechnic Institute

01/2023 - Present

- Managed the club's annual budget, directly contributing to a 100% increase in club attendance.
- Conducted collaborative workshops with industry partners, to help students improve resumes and network

SKILLS

- Languages** - Python (PySpark, Pandas, Matplotlib, Seaborn, Sklearn, Tensorflow, Or-tools), R, SQL, Java, MATLAB
- Analysis Techniques** - Logistic Regressions, Decision Trees, Random Forests, Cosine Similarity, Data Pipelines
- Tools** - VSCode, Databricks, Power BI, Kaggle, Jupyter, PyCharm, RStudio, Azure, Salesforce, Excel, Git