

TAHA ZIAD

Chicago, IL | (734) 660-5563 | taha.ziad.business@gmail.com

Senior Data Analyst with a proven track record of building scalable analytics solutions, automating data pipelines, and validating machine learning models. Expertise in leveraging modern data processing libraries like Polars and Python, developing automated workflows, and integrating large language models to drive operational efficiency. Adept at finding simple solutions to complex problems.

WORK EXPERIENCE

Interactive Brokers, Chicago, IL — Senior Data Analyst

May 2020 - Present

- Engineered high-performance, scalable data pipelines utilizing Polars to extract, process, and validate large-scale SQL tables, supporting machine learning model inputs and outputs.
- Architected a dynamic classification algorithm in R processing millions of records to systematically segment entities for executive-level reporting.
- Accelerated data quality testing for 45+ production models by leveraging statistical analysis to programmatically generate comprehensive unit tests and iterate on discrepancy reviews.
- Designed and deployed a large-scale data validation framework that reduced model review time from months to hours, rigorously tracking feature drift and reverse-engineering logic.
- Conducted comprehensive model validation and data lineage reviews, systematically testing for feature and data drift to ensure ongoing model reliability and accuracy.
- Audited complex systems to identify generation inconsistencies, mitigating system risks and maintaining the integrity of production-level ML architectures.

Housing Information Technology Office, Ann Arbor, MI — Computer Consultant III

October 2015 - May 2019

- Maintained and troubleshooted servers governing access control systems for user facilities.
- Collaborated with backend network infrastructure teams to proactively diagnose and resolve connectivity and access system malfunctions, ensuring uninterrupted network operations.
- Trained and supervised a team of fifteen employees on the intricacies of server infrastructure over a three-year span.

MACHINE LEARNING PROJECTS

Quantitative Algorithm

- Engineered and optimized a quantitative model utilizing reinforcement learning and robust parameter grid search techniques to achieve a 10% performance improvement.
- Implemented strict hyperparameter controls to successfully prevent model overfitting.

Predictive Placement Model

- Developed an end-to-end machine learning pipeline utilizing Python and Jupyter Notebooks to predict competitive outcomes.
- Extracted complex datasets and engineered custom, domain-specific features to significantly improve the predictive accuracy of the final model.

EDUCATION

University of Michigan School of Information, Ann Arbor

Masters of Applied Data Science | Current

University of Michigan School of Information, Ann Arbor

SKILLS & TECHNOLOGIES

Machine Learning & AI: End-to-End ML Pipelines, Advanced Feature Engineering & Selection, Model Selection & Hyperparameter Optimization, Principal Component Analysis (PCA), K-Means Clustering, Reinforcement Learning, Generative AI, Large Language Models (LLMs), Stable Diffusion

Python Data & ML Ecosystem: Scikit-learn, XGBoost, LightGBM, Polars, Pandas, NumPy, Cloudpickle, Joblib, Matplotlib, Seaborn, BeautifulSoup, Pandera, Awpy

Database Architecture & Systems: Relational Database Design, Schema Architecture, Query Optimization, MySQL, AWS, Data Warehousing, Jira (JQL), API Integration (Reddit, Twitter, Spotify)

Languages & Core Tools: Python, R, SQL, Bash/Batch Scripting, VBA, Excel, Tableau, Power BI

R Packages: tidyverse, data.table, ggplot2, lubridate, stringr, MASS, plotly