

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 for risk and compliance use cases. 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

- Advanced from Compliance Analyst to Senior Data Analyst, leading the migration of compliance reporting from Excel to a modern cloud data warehouse to enable scalable, real-time analytics.
- Engineered high-performance data pipelines utilizing Polars to extract, process, and validate large-scale SQL tables for testing machine learning model inputs and outputs.
- Developed a dynamic workforce classification algorithm in R to process millions of internal ticket records, utilizing custom threshold logic to segment employees into operational groups for executive reporting.
- Accelerated data quality testing for 45+ compliance and risk models by leveraging statistical analysis to iterate on discrepancy reviews and programmatically generate comprehensive unit tests.
- Designed a large-scale data validation framework that reduced model review time from months to hours, tracking drift and reverse-engineering logic to ensure regulatory alignment.
- Conducted comprehensive model validation and data lineage reviews, testing for feature and data drift to ensure ongoing model reliability. Audited complex, threshold-based alert systems to identify generation inconsistencies, mitigating risk and maintaining the accuracy of production-level compliance models.

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

October 2015 - May 2019

- Proficiently maintained and troubleshooted servers that governed access control systems for both students and staff within dormitory facilities.
- Collaborated with the backend network infrastructure to proactively diagnose and resolve Wifi connectivity and access system malfunctions, ensuring uninterrupted network operations.
- Demonstrated exceptional leadership skills by effectively training and supervising a team of fifteen employees on the intricacies of the server infrastructure, over a span of three years.

MACHINE LEARNING PROJECTS

Quantitative Trading Algorithm: Engineered and optimized a quantitative trading model for a competitive environment. Utilized reinforcement learning and robust parameter grid search techniques to achieve a 10% performance improvement, implementing 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 race 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