NAWAZISH KHALANDER SHAIK MOHAMMED

Seattle,WA,United States | reachnaazofficial@gmail.com |(315) 952-9825| LinkedIn | Website

PROFESSIONAL SUMMARY

- I am a results-driven Data Analyst with 2+ years of experience delivering user-centered, data-driven solutions across diverse domains. I excel in business analysis, data visualization, and process optimization, leveraging tools like SQL, Power BI, and Tableau to turn complex data into actionable insights. My background includes supporting ETL workflows, ensuring clean and scalable data pipelines, and collaborating with both technical and business teams. I am skilled at creating clear documentation and SOPs, and have successfully implemented AI and machine learning techniques.

EDUCATION

Syracuse University   **Aug 2023 - May 2025**
Master of Science in Applied Data Science, (Excellence Scholarship)

WORK EXPERIENCE

Research Analyst, Syracuse University   **June 2025**
- Built an automated **data pipeline** for a **Language Model**, efficiently harvesting and processing 75,000 books
- Implemented comprehensive text **preprocessing** framework with **custom tokenization** strategies
- **Configured** and **trained transformer**-based **language models** from scratch using HuggingFace Transformers, with custom vocabulary and context length.
- Built a scalable pipeline for training custom language models on 25GB+ text, including data cleaning, tokenization (SentencePiece), and efficient chunking for transformers.
- Optimized model training for limited hardware by applying **quantization** (8-bit), **mixed precision** (fp16), **gradient checkpointing**, and efficient batching strategies.
- Automated **large-scale tokenization and data chunking** (600M+ lines), leveraging memory-mapped arrays and streaming techniques to process datasets exceeding system memory.

Data Analyst, Aptean   **Mar 2022 - Mar 2023**
- Automated **compliance reports** via **Apache Airflow**
- Built 10+ dashboards in **Tableau and Power BI (Power Query, DAX)**, providing real-time insights for business operations
- Documented business processes and created detailed specifications for **data integration** solutions, improving knowledge transfer and system sustainability
- Led root cause analysis sessions with stakeholders to identify and resolve data inconsistencies, implementing preventative measures that improved data reliability
- Maintained data processes using **Excel (Pivot Tables, VLOOKUP, Macros)**, and **Access (Query Design, Forms, Reports)**

Founding Data Engineer, Interaktive XR Solutions   **Jan 2021 - Mar 2022**
- Developed **ML customer risk models**
- Built **CI/CD** workflows using **GitHub** and **Jenkins**, automating **ML pipelines**
- Streamlined **ETL** processes, reducing data processing time by 35% with **Python, SQL**
- Pioneered developing and standardizing **analytical processes** and **BI methodologies**, including interactive **data visualization** dashboards, **exploratory** data **analysis** protocols, and **reporting frameworks**
- Played a **foundational role** in establishing **technical processes** and **methodologies like kanban, agile, waterfall**, **version** control and **code review** practices, which improved team efficiency and code quality

PROJECTS

HnM Fashion Recommendation System| (Deep Learning) (GitHub)
- Engineered a Two-Tower deep learning model integrating multimodal data (text, visuals), enhancing recommendation accuracy by 12%
- Deployed real-time recommendation pipelines with TensorFlow, doubling click-through rates for users

Price Elasticity Modeling and Optimization | Big Data) (GitHub)
- Designed a non-linear Price Elasticity Modeling solution for a shoe manufacturer using PySpark for modeling and SciPy for optimization to determine optimal shoe pricing, increasing sales volume by 1.6%
- Optimized the PySpark model using L-BFGS-B and SLSQP methods in SciPy, increasing revenue by 4%

Predicting Options Price (Data Science and Machine Learning) (GitHub)
- Developed a machine learning model to predict the last price valuation of SPX options using 2023 options chain data, enabling traders to optimize exit strategies and maximize profitability

DASHBOARDS
- Engineered R-Based and Python Based Dashboards (KFIN) , (KDASH)

TECHNICAL SKILLS AND CERTIFICATIONS

- **Programming:** Python, R, Java, Unix, HTML/CSS, C, C++, JavaScript, SQL, Powershell
- **Databases/ ML Engineering Frameworks:** MySQL, PL/SQL, PostgreSQL, SQLite, SSMS TensorFlow, Keras, Scikit-learn, Numpy, PySpark, XGBoost, OpenCV, HuggingFace, RAG, LangChain, PyTorch, Airflow, Apache (Flask, Spark, Kafka)
- **Tools:** Tableau, Power BI, Looker, Git, Jira, Hadoop, MS Excel, MS Access, Streamlit, RStudio, PyCharm, Linux, SharePoint, Power Apps, Docker, Vercel
- **Domain Knowledge:** STEM, Math, Computer Science, Business Growth, Customer Service, Develop Solutions, 3D Solutions