Ritik Toksiya

+1 443-468-1731 — ritiktoksiya99@gmail.com — linkedin.com/in/ritik-toksiya — github.com/Ritiktoksiya99

Summary

Graduate student in Data Science at UMBC with 3 years of experience in data engineering, automation, and business intelligence. Skilled in Python, Azure, Power BI, and SQL for building scalable ETL pipelines and analytics dashboards. Strong foundation in AI, machine learning, and automation workflows with exposure to conversational AI, data pipelines, and healthcare analytics.

Education

University of Maryland, Baltimore County (UMBC) — Baltimore, MD
M.P.S. in Data Science; GPA: 3.883 (Second Semester) — *Aug. 2024 – Present*

KK Wagh Institute of Engineering Education and Research — Nashik, India
B.E. in Electrical Engineering; GPA: 7.9 / 10 — *Jul. 2017 – Jun. 2021*

Experience

Mind Spark Technology — Pune, India
Data Engineer — *June 2021 – June 2024*

- Designed and maintained scalable ETL pipelines using Azure Data Factory and Blob Storage to integrate structured and semi-structured data into SQL Data Warehouse, optimizing schema and performance.
- Developed and deployed Power BI dashboards using DAX, slicers, and drill-throughs for finance and operations teams, automating refresh schedules and implementing row-level security.
- Partnered with BI and finance teams to translate analytical insights into process improvements and cost-saving business decisions.
- Built interactive KPI dashboards in Power BI that reduced manual reporting time by 40% and improved data-driven decisions.
- Created SSIS packages to migrate data from Oracle, flat files, and Excel into SQL Server with automated logging, validation, and error handling.
- Developed SSRS reports (matrix, tabular, and drill-through) to enhance analytics for sales and supply-chain operations.
- Modeled OLAP cubes in SSAS and developed MDX queries to support high-performance reporting and BI dashboards.
- Collaborated with cross-functional stakeholders in Agile sprints using JIRA and implemented CI/CD workflows via Git and Azure DevOps.
- Performed data validation, auditing, and reconciliation across SQL Server, Oracle, and flat files to ensure accuracy and integrity of datasets.
- Executed ad hoc analysis using Excel (Pivot Tables, VLOOKUP, Macros) and SQL to answer stakeholder questions and support data audits.
- Optimized T-SQL queries and Azure Data Factory activities, conducting root-cause analysis to reduce pipeline runtime and improve reliability.
- Documented test cases, QA results, and SOPs to standardize data ingestion, validation, and report generation workflows.

Projects

AI Chatbot for BI Query Automation | *Python, Azure Cognitive Services, LangChain, REST API* — **Aug 2024**

- Developed a conversational chatbot prototype to retrieve Power BI metrics through natural language queries.
- Integrated Azure Cognitive Services and LangChain to manage prompt engineering, RAG pipelines, and real-time metric responses.
- Enhanced accessibility and data interaction for business users by automating Power BI report insights.

Healthcare Claims Anomaly Detection | *Python, Scikit-learn, Pandas, Power BI* — **Jan 2025**

- Applied unsupervised ML algorithms (Isolation Forest, DBSCAN) to identify anomalies in healthcare claim records.
- Created Power BI dashboards for anomaly visualization, data profiling, and audit support.

Technical Skills

Languages: Python, SQL (T-SQL, PL/SQL), Shell Script
AI & ML Tools: Scikit-learn, LangChain, Hugging Face, Azure Cognitive Services, PyTorch
Data & Cloud: Azure (Data Factory, Synapse, Blob Storage), AWS S3, Power BI, Jupyter Notebook
Libraries: pandas, NumPy, Matplotlib, DAX, MDX
Concepts: ETL, Data Warehousing, Data Modeling (Star/Snowflake), Feature Engineering, Model Evaluation, Automation Workflows, Prompt Engineering