Sayed Ashfaq

Bangalore | 302syedashfaq@gmail.com | +91-8074097764 | Linkedin | Github | Scaler

Skills

- **Programming Languages** : Python
- **Libraries and Frameworks** : Flask, Deep Learning, Natural Language Processing (NLP), Fast API, Machine Learning, SVM (Support Vector Machine), EDA
- **Tools** : PyTorch, Tableau, Generative AI, Git, Docker, MySQL, AWS, Google Cloud / GCP, SeaBorn, Google Colab, BigQuery, Kubernetes, REST API, GCP, CI/CD, Prompt Engineering, postman, GitHub

Experience

Associate Analyst, CAMP Systems   Feb 2023 - May 2025
- Processed and analyzed large aviation datasets using **SQL and Python**, delivering insights that supported planning and operational decisions for clients such as **Mountain Aviation and VistaJet**.
- Generated **AI-powered reports** and automated ad-hoc analyses, significantly improving reporting speed and accuracy across teams.
- Designed a **Data Governance POC in Snowflake**, implementing **Row Access Policies, role-based access control, and dynamic masking** for sensitive data (PII).
- Built **Streamlit dashboards** to visualize governance metrics, including top accessed tables and daily query trends, enhancing compliance monitoring.

Highlights: SQL, Python, Snowflake, Streamlit, AI-Powered Reporting, Data Governance, Stakeholder Communication

Projects

RAG_ETE Document Analyzer (Generative AI) | Industry Grade
- Developed an **end-to-end Retrieval-Augmented Generation (RAG) application** enabling users to upload documents and interact with them via conversational AI.
- Integrated **LangChain with FAISS vector database** for efficient semantic search and retrieval of document chunks.
- Implemented **conversation history storage** to maintain context across multiple queries for improved user experience.
- Deployed the application on **AWS**, applying **modular code structure, logging, and exception handling** aligned with MLOps best practices.

Highlights: LangChain, FAISS, Generative AI, RAG, AWS Deployment, MLOps Practices

Customer Churn Prediction using Flask & Machine Learning
- Developed and deployed Customer Churn Prediction models using Random Forest, XBoost, to predict customer churn using real-time inputs.
- Conducted EDA and discovered key insights: high churn in month-to-month contracts and low churn in-term contracts.
- Experimented with different techniques of feature engineering and sampling methods to achieve the 98% accuracy score.
- Deployed models using Flask and Pickle, creating a basic web interface to upload csv file and perform real-time predictions.

Highlights: Machine Learning, Jupyter Notebook, Feature Engineering, Flask, EDA, Deployment.

Data Governance PoC Snowflake + Streamlit (Snowpark)
- Developed Data Governance Proof of concept in Snowflake, implementing row-level access policies.
- Applied dynamic data masking for PII (e.g., emails, SSNs) in Snowflake, ensuring GDPR compliance for 500K+ records.
- Developed a Streamlit app to visualize top accessed tables, daily query trends (excluding weekends), and cron job–based metrics.
- Built a role-based access control model for different user groups and Created process documents and training materials.

Highlights: Snowflake, SQL, Python, Data Governance, Dashboard, usage reporting, Documentation, Streamlit,

Education

Singhania University, Bachelor of Science | 8.2 CGPA   2023

Achievements

- **Google Analytics 4 (GA4) Certificate** — Google Skillshop *(2025)*
- **Data Cleaning Certificate** — Kaggle
- **EDA fundamentals** — Scaler

Course Work - Scaler

Advanced Data Science & Machine Learning Program (2024 – 2025)
Topics: Generative AI, Neural Networks, Deep Learning, MLOps, NLP, ML Algorithms, Project Deployments