

# Sheharyar Akhtar

Lahore, Pakistan | +923015388946 | [Email](#) | [LinkedIn](#) | [Portfolio](#)

## Skills

---

- **Databases:** MySQL, Snowflake, GreenPlum, PostgreSQL, MongoDB, Athena
- **Programming Languages:** Python, R Programming, C++, MATLAB, Stan, SQL, noSQL
- **Big Data & Machine Learning:** Python, Flask/ Dash, LLMs (LangChain, CrewAI, OpenAI)
- **Cloud:** Azure (Machine Learning Studio, Azure OpenAI, Prompt Flow, AI Search), AWS (Glue, Athena, S3, EC2, SageMaker, Lambda, AI Services)
- **Tools:** RStudio, Jupyter Notebook, Anaconda, Git, Docker, MS Excel, MLOps, MLFlow
- **Techniques:** EDA, Causal Inference, Regression, Classification, Clustering, A/B testing, Forecasting, Orchestration
- **Soft Skills:** Agile, Scrum, Leadership, Communication, Problem Solving, Strategic Planning

## Experience

---

### Systems Limited

Aug '23 – Present

#### Consultant – Data Science

- Pioneered uniform model logging using MLFlow across AzureML, Databricks, and On-prem devices: reducing costs through Databricks' complimentary tracking server utility, while enhancing company-wide knowledge transfer efficiency and model assessment.
- Engineered a Python Dash and Snowpark web app on Azure App Service, revolutionizing global ESG data ingestion by automating checks, formatting for Snowflake, and instituting an approval system for production readiness.
- Streamlined Azure DevOps pipeline with Docker for efficient development and deployment, collaborating cross-functionally for standards adherence and user-centric enhancements
- Developed various PoCs surrounding a of data science domains including credit scoring, churn modeling and multi-agent workflows
- Deeply assessed clients' needs based on their vision presented in RFPs to architect solutions, taking budget constraints in mind, on various cloud platforms including Cloudera, Azure, AWS and GCP

### Afiniti

April '21 – Aug'23

#### Data Scientist II | Applied AI

- **Trained and led** up to 5 resources across different accounts across the Iberia region to define strategic roadmaps and ensure high-quality deliverables and timely work.
- Conducted in-depth analyses on clients' environment to understand customer behaviors, which was later used to drive business value and modeling strategies
- Transformed raw data feeds into meaningful modeling datasets using stored procedures in **PostgreSQL and MySQL**
- Developed interactive **RMarkdown/Shiny dashboards** to track and visualize important KPIs.
- **Engineered customer features** that improved model validations by **over 50%** when included in the modeling process.
- Collaborated with an international team of data resources, developers, client executives and project managers to **launch Afiniti on two new clients** in the Iberia region.

## Projects

---

#### Diversity Forecasting Initiative

- Led a Diversity Forecasting Initiative to address gender bias in management, targeting a 45% representation of female Full-Time Employees (FTEs) through predictive models utilizing tree-based algorithms.
  - Engineered interactive simulations allowing exploration of counterfactual scenarios, empowering decision-makers to visualize and assess the impact of strategic changes on female FTE percentages in management.
- Technologies:** Python, lightgbm, scikit-learn, Plotly-dash, Azure (ML Studio, AppService, DevOps), MLFlow, Git, Spark

#### LLM Consultant using Multi-Agent Workflow

- Created a RAG architecture using Azure AI Search to index and vectorize document text and images for retrieval augmented generation
  - Created a multi-agent workflow for assessing RFP requirements, researching relevant literature online and recommending state-of-the-art solutions that satisfy the clients' needs
- Technologies:** CrewAI, Azure OpenAI, Azure AI Search, Azure Machine Learning, LangChain, Google Serper API, custom tools

## Education

---

#### B.S. Mechanical Engineering – GIKI

- **Courses:** Engineering Statistics, Numerical Analysis, Intro to Computing, OOP, Linear Programming, Software Engineering

## Achievements and Awards

---

- Published a Peer-reviewed article in the Cureus Journal of Medical Science, contributing to understanding the root cause of hesitation in getting the Covid Vaccine – Cited 27 times
- Authored multiple articles for the online teaching blog "Towards Data Science," sharing insights on Bayesian modeling – Over 10k views