

MUHAMMED ASLAM

Data Engineer & AI Analyst — Python, SQL, TensorFlow, NLP, MLOps
Calicut, Kerala, India — Open to Relocate or Work Remotely
aslamsadique01@gmail.com — +91 8113841363
LinkedIn — GitHub — Portfolio

Professional Summary

Early-career Data Engineer and AI Analyst with practical experience in building end-to-end data pipelines, predictive modeling, and automation workflows. Demonstrated success in solving real-world problems across churn analysis, fraud detection, and financial forecasting. Skilled in Python, SQL, PySpark, and NLP, with a proven ability to translate complex data into actionable insights for business and risk decision-making.

Technical Skills

Languages: Python, SQL, JavaScript, R, C++

Data Engineering: ETL Pipelines, Web Scraping, REST APIs, Automation Scripts

Big Data & Tools: PySpark (hands-on), Hadoop (basic), GitHub Actions, CI/CD, Docker (basic), Git, Agile

Cloud & Deployment: AWS (EC2, S3 – basic), Firebase, Google Colab, Streamlit

Databases: PostgreSQL, MySQL, MS SQL Server, MongoDB

Analytics: A/B Testing, Forecasting, KPI Dashboards, Time Series Analysis

Visualization: Power BI, Tableau, Matplotlib, Seaborn, Excel

ML/AI: Scikit-learn, TensorFlow, PyTorch, Keras, Transformers, OpenCV

NLP/GenAI: Prompt Engineering, Text Classification, LLM QA, NER, Semantic Search

Statistics: Regression, Hypothesis Testing, Predictive Modeling, Bayesian Inference

Experience

Data Science Intern | Techolas Technologies, Calicut

Jan 2025 – Present

- Developed ML models to predict user churn with 80%+ accuracy using economic features and Scikit-learn.
- Built LLM evaluation pipelines for QA testing, supporting business adoption of generative AI use cases.
- Automated data ingestion and transformation using SQL, Pandas, and Firebase, reducing manual work by 60%.
- Designed Power BI dashboards to visualize retention trends and customer behavior.

Business Analyst Intern | Regalo Pack, Coimbatore

Apr 2025 – Present

- Created SQL-based reporting dashboards for campaign ROI and customer lifecycle analysis in Power BI.
- Unified marketing metrics and established automated workflows for cross-functional teams.
- Delivered insights into pricing optimization, customer segments, and sales funnel performance.

Projects

Healthcare Cost Prediction Tool | Python, Scikit-learn, Power BI

- Built regression models to forecast insurance costs and flagged anomalies using explainability techniques.
- Enabled hospital managers to analyze and plan expenses via interactive Power BI dashboards.

Fraud Pattern Detection Engine (Inspired) | Python, PySpark, Pandas

- Simulated transactional fraud with engineered features like merchant risk scores and time-frequency behavior.
- Applied Spark pipelines for scalable batch analytics and anomaly detection.

RNA Ligand Drug Prediction | Bioinformatics, Streamlit

- Built ML models to predict RNA-ligand interaction affinity; delivered via an interactive web tool for researchers.

JobHunter Automation Pipeline | Web Scraping, Gmail API, LinkedIn API

- Automated job search and outreach using NLP-based resume matching and Gmail/LinkedIn API integrations.

License Plate Detection System | OpenCV, Scikit-learn

- Built contour-based image processing pipeline for license plate detection and OCR-based recognition.
-

Education

Bachelor of Computer Applications (BCA) | IGNOU University, Delhi
Data Science Training Program | Techolas Technologies, Calicut
Digital Marketing Certification | Avodha, Calicut

Ongoing
Jan 2025 – Present
Oct 2024 – Jan 2025

Core Competencies

- Machine Learning, Forecasting, Credit Risk Analytics, A/B Testing
- NLP Pipelines, Prompt Engineering, LLM Evaluation, Generative AI Integration
- SQL Analytics, Big Data (PySpark), Model Validation Monitoring
- Cloud Deployment (AWS, Firebase), CI/CD, GitHub Actions, Workflow Automation
- Agile Reporting, Business Intelligence, Cross-functional Communication

Languages

English (Fluent – C1) — Hindi (Fluent) — Malayalam (Native) — Urdu (Advanced)