MOHAMMED FAIZAAN

+1 (562) 583 5550 | Portfolio | GitHub | faizaanmd6090@gmail.com

Education

MS: Computer Science **2025-2027**
California State University Long Beach
B.E: Computer Science and Engineering **2020-2024**
Navodaya Institute of Technology, Raichur (Affiliated to VTU)

Experience

Compsoft Technologies | Intern **Mar 2023 – Apr 2023**

- Researched predictive modeling techniques such as **logistic regression** and **gradient boosting** for consumer credit risk estimation, analyzing payment history, utilization, and delinquency trends.
- Engineered key behavioral features (rolling averages, utilization ratios, payment delays) and evaluated models using ROC-AUC, KS statistic, and calibration plots.
- Built **Python + Spark** workflows for model retraining, validation, and drift detection, ensuring consistent performance across time-based cross-validation splits.
- Presented findings to faculty and industry partners with Power BI dashboards and executive summaries, demonstrating clear communication and analytical rigor.

Cranes Varsity | Intern **Aug 2023 – Sep 2023**

- Built **RESTful APIs (Node.js + PostgreSQL)** for transaction tracking and fraud detection services.
- Implemented **unit/integration testing (JUnit, Jest)** improving code reliability and reducing bugs in production by **35%**.
- Integrated **Kafka-based event streaming** for processing 500K+ financial events/day in near real-time.
- Worked in an **Agile team**, participating in sprint planning, peer code reviews, and weekly demos with stakeholders.

Kodnest | Intern **Mar 2024 – May 2024**

- Collaborated with a cross-functional team to design and deploy **machine learning models** for forecasting, anomaly detection, and classification on large-scale enterprise datasets (structured + unstructured).
- Built and optimized **end-to-end data pipelines** using **Python, SQL, and Pandas**, improving data processing efficiency by **28%**.
- Contributed to **LLM-based agent development**, including **prompt engineering, tool integration, and fallback orchestration**, enabling automated insights for business operations.
- Applied **feature engineering** and **model interpretability techniques**, boosting model accuracy by **15%** while ensuring compliance with explainability standards.
- Assisted in deploying models to **cloud environments (AWS S3, Lambda, SageMaker)**, reducing inference latency by **20%** and supporting scalable enterprise usage.

Projects

Data Reliability & BI Automation Platform

- Implemented an end-to-end data reliability platform to centralize ETL, validation,exception reporting for cross-departmental business datasets.
- Built automated ETL pipelines **(Python + SQL)** to ingest payroll, sales, and inventory CSVs/SQL extracts; standardized schemas and implemented column-level cleansing rules.
- Implemented comprehensive data integrity checks **(completeness, uniqueness, schema, range, referential)** that ran as scheduled Airflow jobs and logged results to a monitoring table.
- Developed scalable exception reporting: automatically generated summary reports and **CSV** error extracts; emailed actionable owner-specific alerts with remediation steps for each exception.
- Created an interactive **Power BI** workspace with dashboards showing data freshness, error rates by source, exception backlog, and business KPIs (revenue by product, inventory health).

World Happiness Index Analysis [Data Science] | ***[GitHub]***

- Conducted ***exploratory data analysis (EDA)*** on dataset to uncover correlations between GDP, social support, and happiness scores.
- Utilized Python libraries such as ***Pandas*** and ***NumPy*** for data wrangling, cleaning, and statistical analysis.
- Built visualizations using ***Matplotlib/Seaborn*** and performed analysis, improving prediction ***accuracy by 10%*** on cleaned datasets.
- Identified patterns and correlations to explain the influence of socioeconomic factors on national well-being.

Customer Segmentation & Offer Propensity (Simulation)

- Performed **PCA + K-Means** on standardized cardholder spend and merchant-mix features (MCC share, recency/frequency/monetary), tuning **k** via silhouette and inertia to derive stable, actionable segments.
- Built a lightweight **uplift model** (Two-Model and **XLearner**) with propensity adjustment; evaluated using **Qini/uplift AUC** and calibrated treatment thresholds for targeting.
- Ran **simulation / A/B replays** on historical campaigns with time-based splits; estimated **~20% predicted lift** vs. baseline targeting and computed confidence intervals via bootstrap.
- Packaged a **reproducible pipeline** (Python/Spark notebooks) with feature dictionaries, data-quality checks, and **MLflow (local)** experiment tracking for consistent re-runs and audits.

Skills & abilities

Languages: Python, C++, C, Java, SQL, HTML, CSS, JavaScript, SQL
AI/ML: Retrieval-Augmented Generation (RAG), LLMs, Hugging Face Transformers, NLP, FAISS, Computer Vision, Deep Neural Networks
Libraries/Frameworks: NumPy, Pandas, Matplotlib, Seaborn, PyTorch, TensorFlow (Keras), Scikit-learn, ReactJS
Cloud & DevOps: AWS (EC2, EBS, IAM), Git, Google Colab, Docker, Linux

Leadership and Certifications

- Attended Honeywell-hosted cybersecurity workshop covering threat modeling and ethical hacking basics.
- Led a 6-member team to ***1st place*** in a hackathon by completing 4 timed challenges in algorithms, debugging & system design.
- Completed ***"Convolutional Neural Networks in Python"*** on Udemy, implementing image classifiers using Keras.
- Captained a 50-member team to win a major college event through strategic coordination and leadership.