PRATHYUSHA REDDY ALLAM

Buffalo, NY | +1-716-910-9655 | allamprathyusha77@gmail.com

PROFESSIONAL SUMMARY

Computer Science graduate student specializing in Artificial Intelligence, Machine Learning, and Data Analytics with 2+ years of industry experience at Tata Elxsi. Skilled in designing end-to-end data pipelines, automating reporting workflows, and delivering actionable insights from large-scale datasets using Python, SQL, and cloud platforms. Proven ability to monitor KPIs, perform root cause analysis, and communicate findings to cross-functional stakeholders. Published IEEE researcher with hands-on project experience in fraud detection, customer segmentation, and sales analytics.

TECHNICAL SKILLS

Languages: Python, SQL, C++, C
AI / ML: TensorFlow, PyTorch, Scikit-Learn, Deep Learning, Computer Vision, NLP, CNN, LSTM, Transformers
Data Engineering: Pandas, NumPy, Feature Engineering, Data Preprocessing, ETL Workflows, Data Cleaning & Validation
Cloud & Tools: Microsoft Azure, Git, Docker, Linux, Jupyter Notebook, PyCharm, Power BI, Tableau
Analytics: Statistical Analysis, KPI Reporting, Root Cause Analysis (RCA), A/B Testing, Business Intelligence

PROFESSIONAL EXPERIENCE

Engineer — Tata Elxsi | India *Aug 2022 – Jun 2024*

- **Analysed large-scale operational datasets** (10M+ records) using Python (Pandas, NumPy) to identify performance trends, system anomalies, and behavioural patterns — enabling data-driven engineering and business decisions.
- **Extracted, cleaned, and transformed data** from Azure cloud environments and internal databases to support validation, comparative reporting, and downstream analytics workloads.
- **Monitored and validated Key Performance Indicators** (KPIs) from telemetry and Katana data sources, producing recurring performance reports that improved system reliability tracking and operational efficiency visibility.
- **Conducted structured root cause analysis** (RCA) on critical system events by correlating data across multiple sources, identifying key failure drivers, and presenting mitigation findings to engineering stakeholders.
- **Automated data extraction, preprocessing, and reporting pipelines** using Python scripts, reducing manual analysis time by ~40% and improving reporting accuracy and consistency.
- **Built interactive Power BI dashboards** and technical analytical reports, presenting actionable insights and recommendations to both engineering teams and business leadership.

Data Engineering Intern — Tata Elxsi | India *Feb 2021 – Mar 2022*

- **Collected, validated, and analysed project datasets** to support research and engineering activities, ensuring data completeness and accuracy before downstream processing.
- **Performed structured data cleaning and quality checks** across multiple data sources, reducing data inconsistencies and improving report reliability for project stakeholders.
- **Assisted in generating analytical reports** and progress summaries to support project tracking, milestone reviews, and cross-team decision-making.
- **Maintained organized datasets**, technical documentation, and version-controlled project records using Git — facilitating knowledge transfer and team collaboration.
- **Collaborated with senior engineers** to interpret data trends and present findings during weekly project reviews, strengthening communication of technical insights to non-technical audiences.

KEY PROJECTS

Credit Card Fraud Detection — Python*, Scikit-Learn, XGBoost, Pandas, ROC-AUC*

- **Built and evaluated multiple classification models** (Logistic Regression, Random Forest, XGBoost) on a highly imbalanced credit card transaction dataset, applying SMOTE oversampling and feature engineering to improve minority class detection.
- **Optimized model performance** using precision, recall, F1-score, and ROC-AUC metrics — achieving a 92%+ AUC score — reducing false negatives critical to fraud prevention systems.
- **Conducted feature importance analysis** to identify the top predictors of fraudulent behavior, enabling interpretable outputs for business and compliance stakeholders.

Sales Performance Analytics Dashboard — SQL*, Python, Tableau / Power BI*

- **Developed an interactive multi-page dashboard** tracking sales KPIs — including revenue trends, product performance, regional breakdowns, and year-over-year growth — across multiple business segments.

- **Built automated ETL workflows** using Python and SQL to extract, clean, and transform raw sales data into structured analytical datasets, reducing manual data preparation time by ~35%.
- **Delivered drill-down visualizations** enabling sales and operations leadership to self-serve on performance metrics, replacing ad-hoc Excel-based reporting.

E-Commerce Customer Segmentation — Python*, K-Means, RFM Analysis, Tableau*

- **Applied K-Means and hierarchical clustering** on RFM (Recency, Frequency, Monetary) features to segment e-commerce customers into behavioural cohorts — identifying high-value, at-risk, and new customer groups.
- **Engineered customer-level aggregate features** from raw transaction logs using Python, transforming unstructured purchase history into a clean, analysis-ready dataset.
- **Created visual reports and Tableau dashboards** presenting segment profiles, enabling the marketing team to design targeted retention and acquisition strategies with measurable business impact.

End-to-End ETL Data Pipeline for Retail Analytics — Python*, PostgreSQL, Azure Blob Storage, Power BI, Pandas*

- **Designed and implemented a fully automated ETL pipeline** ingesting raw retail transaction data from multiple CSV and API sources into a normalized PostgreSQL data warehouse, processing 500K+ records per batch with error logging and retry logic.
- **Applied data cleaning, deduplication, and schema validation routines** using Python (Pandas) to standardize inconsistent data formats across sources — improving downstream data quality and reducing null rates by ~45%.
- **Staged intermediate datasets in Azure Blob Storage** and orchestrated batch transformation jobs using Python scripts, enabling scheduled daily refreshes without manual intervention.
- **Built a Power BI reporting layer** on top of the warehouse, delivering automated dashboards tracking product inventory turnover, sales velocity, and regional demand trends — replacing manual weekly Excel reports.

EDUCATION

M.S. Computer Science & Engineering (AI & ML) — State University of New York at Buffalo | Buffalo, NY*Aug 2024 – Dec 2025*

Relevant Coursework: Machine Learning, Deep Learning, Natural Language Processing, Computer Vision, Data Analytics, Cloud Computing, Algorithm Design

B.Tech, Computer Science & Engineering — Karunya Institute of Technology and Sciences | India*, Jun 2018 – May 2022*

CERTIFICATIONS & PUBLICATIONS

Publication: Allam, Prathyusha Reddy. "Hand Gesture Recognition using Shape-based Image Features for Music Controller." CISES 2022 — IEEE Xplore.