MONISHA PATRO

+1-812-345-4652 monishaapatro@gmail.com linkedin.com/in/monisha-patro/ GitHub Portfolio

EDUCATION

INDIANA UNIVERSITY BLOOMINGTON — **United States**
Master of Science in Data Science — **August 2023 – May 2025**
VELLORE INSTITUTE OF TECHNOLOGY — **India**
Bachelor of Technology in Computer Science — **June 2019 – May 2023**

WORK EXPERIENCE

Therapprove | ***Intelligent care-matching for pediatric therapy access*** — **June 2025 - Present**
Data Science Intern
- Developed Machine Learning models to match children with optimal pediatric therapists by leveraging caregiver input, provider availability and therapy specialization reducing referral-to-response time from weeks to near immediate.
- Analyzed CRM and referral queue metadata using Postgres to identify client bottlenecks by therapy type, location, and insurance coverage surfacing trends that reduced average intake wait time and informed queue prioritization strategies.

Candid | ***Nonprofit intelligence via data integration***
Data Science Intern — **May 2024 – December 2024**
- Developed and implemented scalable SQL-based ETL processes within the data warehouse platform to process and standardize non-profit data from govt. publications (~10M+ records), accelerating data delivery by 25% for analytics and product teams.
- Partnered with Data Services and API engineering teams to integrate cleaned and mastered datasets into public-facing APIs as a data as service, enhancing data accessibility for 10K+ external users while maintaining backward compatibility using SQL.
- Collaborated with cross-functional product & engineering teams to translate stakeholder requirements into Power BI dashboards, informing product development and strategy, contributing to a 15% increase in product adoption while tracking key KPIs.

EProtons | ***Real – Time analytics for EV stations*** — **February 2023 – July 2023**
Data Science Intern
- Reconfigured PostgreSQL indexing strategies to improve query performance by 27%, on energy datasets used in forecasting models.
- Engineered distributed data pipelines on AWS EMR and PySpark using Python, orchestrating parallel data processing workflows, yielding a 5x acceleration in large-scale analytics tasks.
- Designed and evaluated an A/B test comparing flat-rate and dynamic pricing models across EV charging stations, informing decisions based on KPIs utilizing data models, uncovering a 12% lift in session completion using SQL to control for location-based confounders.

Mukham | ***AI – driven attendance and geofencing*** — **October 2022 – March 2023**
Data Analyst
- Spearheaded development of CNN-based facial authentication models, cutting spoofing incidents by 50% across security endpoints.
- Augmented fraud detection performance utilizing geolocation & time-series signals into predictive models, increasing precision by 35%.
- Established image processing pipeline for facial data, improving image quality for 80% of enrolled users and minimizing the number of support tickets related to image failures.

PROJECTS

eBay Product Strategy Analytics
- Engineered a dataset of items by integrating eBay's Browse & Marketplace APIs, combining product metadata with sales to support marketplace performance analysis; cleaned & transformed NDJSON into structured CSV and Parquet files for scalable analytics.
- Interpreted impact of listing strategies like auction vs fixed-price & single vs multi-variation formats, finding multi-variation listings had 12% higher sell-through rates while auctions underperformed for low-demand items, informing pricing & listing recommendations.

TelConnect Customer Churn Prediction
- Designed a machine learning pipeline on 1M+ customer records to predict churn, using AdaBoost achieving high recall (81%) and improving early identification of at-risk users.
- Generated real-time churn scores and cohort reports segmented by tenure, complaints, and service usage, uncovering key drop-off patterns and enabling targeted retention strategies through a stakeholder-facing dashboard.

Amazon Product Review Analysis
- Fine-tuned Machine Learning models on 1M+ Amazon product reviews to classify sentiment with 91% accuracy, uncovering how language tone, category, and phrasing patterns reflect customer experience across product types.
- Refined preprocessing with advanced tokenization, attention masking, and truncation logic to handle domain-specific slang and variable review lengths & explored misclassification patterns in sarcastic/mixed-tone reviews to boost robustness in sentiment detection.

SKILLS

- **Programming & Databases**: Python, SQL, R, PySpark, ETL.
- **Statistics & Experimentation**: A/B Testing, Causal Inference, Hypothesis Testing, Forecasting, Correlation Analysis, Time-series Analytics, Predictive & Descriptive Modeling, Experimentation.
- **ML & Visualization**: Power BI, Tableau, Regression, Classification, Gradient Boosting, Bagging, SVM, TensorFlow, PyTorch.