Srujani Mareddy

Atlanta, GA, 30324 | +1 708-769-5970 | srujanimareddy@gmail.com | LinkedIn | GitHub

EDUCATION

MS in Data Science and Analytics - Georgia State University, J. Mack Robinson College of Business (GPA: 4.0) *December 2024*

BTech Electronic & Communication Engineering – Malla Reddy Engineering College (GPA: 3.5) *June 2021*

TECHNICAL SKILLS

Programming & Data Tools: Python, R, SQL, SAS, Bash, PySpark, SparkSQL

Data Engineering: ETL Testing, Data Validation, Data Preprocessing, Feature Engineering, Data Profiling, Statistical Hypothesis Testing

Machine Learning & Advanced Analytics: Supervised & Unsupervised Learning (Classification, Clustering, Regression), Ensemble Learning, Deep Learning, Neural Networks, Predictive Analytics, A/B Testing, Natural Language Processing, LLM, Langchain, Hugging Face

Cloud Platforms, Data Visualization & BI Tools: Azure SQL Database, Azure Data Factory, ADLS Gen2, Azure Synapse Analytics, Databricks, Power BI, Matplotlib, Seaborn, Excel, Tableau, Alteryx, Qlik Sense

Business Analytics: Text Analytics, Predictive Modeling, KPI Tracking, Quantitative Analysis, Statistical Modeling, Time-Series Forecasting

Certifications: **Databricks Certified Data Engineer Associate**, **Microsoft Certified Power BI Data Analyst Associate**

WORK EXPERIENCE

Sales and Product Analyst– MD Direct Supply, Remote *January 2025 – Present*

- Automated supplier price comparisons and reduced manual effort by **30%** by developing an Excel-based price aggregation tool using Power Query, VLOOKUP, and VBA macros, consolidating multi-supplier data into a structured format for faster decision-making.
- Enhanced cost analysis for 15+ high-demand products by implementing **SQL-based data extraction**, **Excel Pivot Tables**, and Python for **price fluctuation tracking**, which identified negotiation opportunities through **variance analysis and moving averages** and recommended supplier shifts with a potential **2-5%** cost savings on bulk orders.

Data Analyst Intern – MidPay, Atlanta, GA, USA *May 2024 – August 2024*

- Built a real-time fraud monitoring dashboard in **Power BI** with **SQL**, tracking high-risk payments, failed attempts, and geolocation mismatches, reducing fraud response time.
- **Developed a user risk segmentation model** using **K-Means clustering and rule-based classification**, flagging **50+ high-risk users** and preventing **50+ fraud-related disputes in three months**, doubling fraud detection accuracy**.**

Research Data Analyst, Truist Bank (GSU), Atlanta, GA, USA *August 2024 – December 2024*

- Built NLP pipelines using OCR, tokenization, TF-IDF, and embeddings for data extraction, leveraging advanced generative models like BERT, GPT-3.5, and Llama 2, which delivered actionable insights from complex financial text data.
- **ESG Data Benchmarking:** Partnered with Truist Bank's ESG Data Analytics team to create data-driven industry benchmarks, improving the accuracy of climate-related financial disclosures by **15%** and facilitating better ESG compliance.

Research Data Analyst, Southeast Regional Cooperative (GSU), Atlanta, GA, USA *August 2023 – August 2024*

- Reduced weekly fresh produce waste by **12+ tons** using an **ARIMA-based time series forecasting model** trained on **24 months** of demand data, optimizing supply allocation and minimizing spoilage.
- Increased produce shelf life by **3-5 days per cycle** by implementing a **First-Expiration-First-Out** inventory system integrated with a real-time **Power BI dashboard**, enabling automated surplus tracking, expiration-based prioritization, and dynamic redistribution alerts to minimize spoilage and optimize food bank distribution.

Data Analyst, Infosys Limited (Client: Myntra**), Hyderabad, India** *April 2022 - July 2023*

- Increased seller **Net Promoter Score** by **33%** by developing a real-time **Seller Reimbursement Funnel** Dashboard in Power BI, integrating SQL-based ETL pipelines to track claims, approvals, and return trends.
- Prevented **$500,000** in losses and reduced fraudulent payouts by analyzing return transactions, identifying **high-risk sellers,** and implementing an automated **risk scoring model** in SQL & Python (Logistic Regression, Anomaly Detection, XGBoost) to flag suspicious transactions before reimbursement approval.
- Cut fraudulent reimbursements and quality check costs by **15%** by analyzing orders using **SQL queries and Python-based data modeling**, identifying sellers with higher claim rates, and optimizing quality control interventions.
- Improved payout accuracy to **93%** by building a **predictive reimbursement model** using XGBoost, Random Forest, optimizing financial planning and reducing discrepancies.

Data Engineer, CtrlS Datacenters (Client: Care Hospitals**), Hyderabad, India** *April 2021 –April 2022*

- Designed and implemented an end-to-end data pipeline using **Azure Data Factory, Synapse, and Spark** to ingest and transform patient records, and billing from 8 hospital branches, reducing query execution time.
- Developed and maintained **Power BI dashboards** tracking **ICU bed utilization**, emergency department wait times, and patient discharge trends, reducing ICU congestion and improving bed turnover time. Provided real-time insights that enabled hospital administrators to optimize staffing and resource allocation.
- Implemented a predictive readmission risk model using Azure Machine Learning and analyzing patient records to reduce unplanned readmissions by **14%**. Enabled proactive patient interventions, improving recovery outcomes and optimizing ICU bed allocation.

PROJECTS

AIRA 1.0 – AI-Powered RAG: Engineered a Gen AI-based RAG system with FAISS & GPT-3.5, enabling fast document retrieval & query answering.

Walmart Sales Forecasting: Built a Random Forest & Exponential Smoothing model on Azure ML, improving forecasting accuracy by 90%.

Sleep Sphere Analytics: Visualized health data insights using Google BigQuery & Looker Studio, enhancing data-driven decision-making.

Telecom Churn Prediction: Predicted customer churn with 89% accuracy using ML models, optimizing retention strategies via Power BI dashboards.