Aditi Reddy Doma

aditidoma@gmail.com | 917-337-7045 | linkedin.com/in/aditireddydoma | GitHub | Portfolio

EDUCATION

Masters of Science in Data Science — ***Expected: December 2026***
Rochester Institute of Technology (*GPA: 4.0*) — *Rochester, New York*

Bachelor of Technology in Artificial Intelligence — ***July 2020 - May 2024***
Mahindra University (*GPA: 3.2*) — *Hyderabad, India*

WORK EXPERIENCE

Data Engineer Intern — ***May 2025 - Present***
icitizen — *Remote, USA*

- Extracted and analyzed **500k+ unstructured congressional profiles** from Congress.gov and Bioguide into **structured datasets** using **AWS (DynamoDB, S3, Glue, Lambda)**, enabling real-time updates on political representation and funding interests.
- Reduced **research time by 23%** and enhanced policy engagement by designing interactive U.S. maps in **QuickSight**, linking states to representatives and senators, highlighting their sponsored bills and funding interests for student government advocacy.

Graduate Teaching Assistant — ***January 2025 - Present***
Rochester Institute of Technology — *Rochester, New York*

- Guided 200 undergraduates in **database systems** and **SQL**, through labs focused on query writing, data normalization, and relational modeling using industry-aligned case studies.

Business Analyst Intern — ***December 2023 - July 2024***
StandardWings Technologies Pvt. Ltd. — *Nashik, India*

- Documented client requirements by collaborating with **15+ cross-functional** team members and **attending client meetings**.
- Extracted **actionable insights** from **20,000+ rows** of patient data using **MS Excel**, then designed a low-code application in **Mendix**, leading to improved patient onboarding and streamlined appointment scheduling.
- Presented the Mendix app functionality to technical and non-technical stakeholders using clear **MS PowerPoint** decks.

Machine Learning & Data Analyst Intern — ***June 2023 - August 2023***
Cloud4C — *Hyderabad, India*

- Achieved **Natural Language Processing (NLP) accuracy above 90%** by developing a sentiment analysis product using **BERT** through transfer learning and fine-tuning on customer raised tickets.
- Optimized **TensorFlow** training pipelines on **Databricks (PySpark, SQL)** using adaptive learning rates for large-scale text analysis.
- Identified the **top 5** recurring customer concerns by applying **K-Means clustering** on BERT embeddings.

PROJECTS

Portfolio Allocation Model (GitHub)

- Researched and scraped **5 years of stock data** for 9 companies across 3 sectors using Yahoo Finance, and optimized portfolio allocation through **mean-variance modeling** using **Pyomo, NumPy, and Pandas**.
- Identified Buy & Hold as the optimal investment strategy with **15% ROI** by performing **statistical analysis** of risk-return trade-off assessments using **Monte Carlo simulation** in **Python**.

Strategic Insights for Mitron Bank's New Credit Card Line (GitHub)

- Developed **3 Power BI dashboards** on income, expenditure trends, customer segmentation with **Power Query** and **visuals**.

Energy Consumption Predictor (GitHub)

- Forecasted hourly energy consumption with **92.54% accuracy** using **time-series Linear Regression** in **R** for 5000+ homes.
- Built a **Shiny dashboard** to show how temperature, time and house attributes **influence future energy consumption**.
- Visualized county level energy hotspots using **ggplot2** and **plotly**, highlighting high consumption zones for energy insights.

Covid Data Exploration and Visualization (GitHub)

- Cleaned data using **Excel**, automated updates via **Power Automate** and conducted **exploratory analysis** with **SQL queries**.
- Showcased COVID-19 trends and vaccination impact in **Tableau** using **multiple visualizations** including charts, maps.

AutoConvo: Real-Time Driver Intent Detection (GitHub)

- Designed a speech-to-intent pipeline using **Whisper ASR** and **DistilBERT** to classify 40+ in-car voice commands at **92% accuracy**.
- Integrated context-aware dialogue handling with dynamic fallback and exposed predictions via lightweight APIs for in-car apps.
- Optimized deployment on Jetson Nano with **TensorRT** and **CUDA**, enabling real-time inference under 400ms latency.

IOT Telemetry Analytics for Health Monitoring Devices (GitHub)

- Analyzed **2M+ IoT telemetry time-series records** simulating health devices (glucose, heart-rate, temperature sensors).
- Reduced **false alarms by 18%** using a **Kafka + PySpark** pipeline and **anomaly detection models (Isolation Forest, ARIMA)**.

SKILLS

- **Programming & Libraries:** Python, Java, R, MySQL, Pandas, NumPy, Scikit-Learn, Matlab, Keras, TensorFlow, PyTorch, Git
- **Data Tools:** Tableau, PowerBI, Seaborn, Excel, Word, PowerPoint, Databricks, AWS (S3, Redshift, Lambda), Kafka, Apache Spark
- **Machine Learning:** Supervised & Unsupervised Learning, NLP, A/B Testing, Predictive Modeling, Statistical Analysis