

Rahul Teja Bolloju

📞 906-299-2760 — ✉ b.rahulteja01@gmail.com — 🏠 Chicago, IL — 🌐 in/rahultejabolloju — 📁 portfolio

Summary — Results-driven **Data Scientist** with **3+ years of experience** in machine learning, time series forecasting, and data visualization. Proficient in **Python, SQL, Power BI, and Azure ecosystem (Azure ML, Data Factory, Fabric, Cognitive Services)**. Proven ability to build end-to-end analytical solutions, automate data workflows, and drive business decisions through actionable insights.

Education

Michigan Technological UniversityApril 2024

Master of Science in Data Science, Houghton MI

Courses: Operations Management, Info Systems Management, Statistical Methods, Time Series Forecasting, Business Analytics

Skills

Python	SQL	R	Git
Azure ML	Azure Data Factory	Microsoft Fabric	Cognitive Services
Machine Learning	Time Series Forecasting	Deep Learning	Statistical Modeling
Power BI	Data Visualization	Forecast Evaluation	Data Storytelling
Problem Solving	Cross-Validation	Feature Engineering	Leadership

Experience

Pitney BowesStamford CT

Data Scientist (SendPro C Lite)August 2023 – Present

- Reduced customer churn by 10% using machine learning models (Random Forest, XGBoost) in Python and Azure Machine Learning, achieving an AUC-ROC score of 0.87.
- Extracted and cleaned data with SQL and Azure Data Factory, implementing ARIMA and Holt-Winters models to forecast churn, reducing attrition by 15% through targeted prevention strategies.
- Visualized model performance using ROC curve and precision-recall charts, sharing insights via interactive dashboards in Azure Power BI, facilitating clear communication to stakeholders.
- Optimized pricing strategies with Bayesian optimization and hyperparameter tuning in Azure ML, improving model accuracy by 20% and increasing customer lifetime value by 10%.

Data Science Intern (SendTech)July 2023 – April 2024

- Predicted defaulters using XGBoost in Python, achieving 0.76 AUC to improve high-risk customer outreach.
- Designed and deployed Power BI dashboards with Microsoft Fabric and Azure, improving debt visualization and identifying actionable KPIs.
- Unified and analyzed multi-source data with SQL, Azure Data Factory, and Microsoft Fabric Lakehouse, processing 12M+ records to optimize debt collection and improve agent productivity.
- Refined ARIMA and Prophet models in Python for time series forecasting, reducing MAPE by 15% for 60% of clients with advanced cross-validation.

FonkR SolutionsHyderabad, India

Data ScientistJune 2020 – July 2022

- Extracted and visualized purchasing trends through web scraping with Python (Beautiful Soup and SQL, building interactive dashboards in Power BI and Plotly to drive strategic decisions.
- Implemented sentiment analysis using Azure Cognitive Services and ML models (SVM, Logistic Regression), achieving 88% accuracy and improving customer experience by 20%.
- Applied agile methodologies to enhance data quality, achieving a 20% improvement in accuracy and a 15% reduction in errors via iterative process optimization and data governance best practices.

Projects

Spondylitis ResearchDecember 2023 - Present

- Working with a surgeon to analyze a dataset of 1,500 patients, identifying key factors in spondylitis using machine learning models (Random Forest, AdaBoost). Achieved 52% accuracy with AdaBoost while addressing data issues.

Real-Time Epileptic Seizure Prediction with Deep Neural NetworksAugust 2023 - April 2024

- Utilized Microsoft Fabric for real-time data streaming and deployment of deep learning models, achieving a 98.28% accuracy in predicting epileptic seizures from EEG data.

Chipotle's Supply Chain Management with RFIDJanuary 2024 - April 2024

- Conducted supply chain analysis for Chipotle, assessing vendor logistics, workflow efficiency, and cost trade-offs; recommended RFID adoption to improve inventory accuracy, reduce waste, and enhance delivery performance.

Houghton Gas Station Statistical AnalysisJanuary 2024 - April 2024

- Led a team in data collection at Houghton Gas Station to identify workflow bottlenecks, used Two-way ANOVA in R to analyze peak-hour congestion, reducing wait times by 20% through process optimizations.