

# SOWMYA RAMISETTY

Hartford , CT | +1 (603)-451-0629 | [ramisettysowmyadata@gmail.com](mailto:ramisettysowmyadata@gmail.com) | [LinkedIn](#) | [Portfolio](#)

## PROFESSIONAL SUMMARY

Data Scientist with 2 years of experience applying machine learning, statistical modeling, and data visualization to deliver business insights and scalable AI solutions. Proficient in Python, SQL, TensorFlow, and Power BI with hands-on expertise in predictive modeling, NLP, and big data processing using Spark. Skilled at building and deploying end-to-end ML pipelines on cloud platforms (AWS, Azure), automating workflows, and collaborating in Agile teams to drive data-driven decision-making and measurable business outcomes.

## TECHNICAL SKILLS

**Programming & Querying:** Python, R, SQL

**Machine Learning & AI:** Scikit-learn, TensorFlow, PyTorch, XGBoost, NLP (spaCy, Hugging Face Transformers)

**Data Visualization:** Tableau, Power BI, Matplotlib, Seaborn, Plotly

**Big Data & Cloud:** Apache Spark, AWS (S3, EC2, SageMaker), Azure Machine Learning

**Databases:** MySQL, PostgreSQL, MongoDB

**MLOps & Workflow Automation:** MLflow, Airflow, Docker, Kubernetes, Git

**Core Competencies:** Data Wrangling, Feature Engineering, Statistical Modeling, Predictive Analytics, Model Deployment

## PROFESSIONAL EXPERIENCE

### Data Scientist

Dec 2021 - Dec 2023

#### Accenture Ltd | Bangalore, India

- Designed predictive models in Python and XGBoost that improved client sales forecast accuracy by 15%, helping business units allocate resources more effectively.
- Reduced ETL bottlenecks by 40% through rebuilding pipelines with SQL and Apache Spark, which provided near real-time data access for analytics teams.
- Automated feature engineering and preprocessing in Pandas/NumPy, saving analysts over 10 hours each week and accelerating delivery of machine learning prototypes.
- Uncovered key customer sentiment drivers by applying NLP with spaCy and NLTK to 200K+ feedback records, contributing to an 18% boost in service ratings.
- Built interactive Tableau dashboards that cut reporting turnaround time by 25%, allowing executives to monitor KPIs with live drill-down insights.
- Identified customer micro-segments through K-Means clustering, which powered targeted campaigns and drove 12% higher engagement rates in pilot regions.
- Guided product teams with statistical testing and A/B experiments, ensuring new feature rollouts delivered measurable ROI improvements before scaling.
- Collaborated in Agile sprints, integrating data science models into cloud-hosted client platforms on AWS, which improved adoption and reduced deployment delays.

## PROJECTS

### Hospital readmission risk prediction

- Ingested and normalized 3M+ EHR and claims records using SQL and Apache Spark, engineered 120 clinical and temporal features with Pandas to improve signal coverage and reduce missing-rate by 28%.
- Trained ensemble models with XGBoost and LightGBM using stratified cross-validation and Optuna tuning, achieving ROC-AUC 0.89 and a 2.3x lift versus baseline logistic regression.
- Packaged models with MLflow and Docker and deployed as a SageMaker REST endpoint, integrated into Power BI dashboards to enable clinician targeting that reduced 30-day readmission by 12% in a controlled pilot.

### Customer churn & retention modeling

- Aggregated two years of transactional and engagement logs via SQL and Spark, engineered RFM, sequence and behavioral features in Pandas, increasing predictive feature coverage by 28%.
- Built a stacked LightGBM + TensorFlow model, applied SHAP explainability and recursive feature elimination, producing AUC 0.87 and improving precision for high-risk customers by 18%.
- Operationalized daily scoring with Airflow, Docker, and MLflow on AWS ECS, enabling targeted retention campaigns that lowered monthly churn by 9% and lifted customer LTV by 7%.

## EDUCATION

### Master of Computing and Information Systems

Jan 2024 - Oct 2025

New England College | Henniker, New Hampshire

### Bachelor of Technology in Computer Science Engineering

Jun 2017 - Oct 2021

QIS College of Engineering and Technology | Ongole, India

## CERTIFICATIONS

- Generative AI with Large Language Models - **Coursera**
- SQL for Data Science - **LinkedIn Learning**
- IBM Data Science Professional Certificate - **Coursera**
- Applied Data Science with Python Specialization - **Coursera**