

# Ashritha Gugire

[gugireashritha3@gmail.com](mailto:gugireashritha3@gmail.com) • 949-204-6952 • Fairfax, VA • LinkedIn • Website

## SUMMARY

Data Analyst Since 2020 and with 4+ years of experience in Data Science, Machine Learning, Research and Development. Proficient in Python, SQL, building ML pipelines with developing interactive dashboards using Tableau, Excel, and custom scripts. Expertise in automate algorithms, Predictive Modeling, Validation, working with both structured and unstructured data for data-driven decision-making. Experienced with cloud platforms and communicating effectively to both technical and non-technical audiences.

## TECHNICAL SKILLS

**Programming:** Python (NumPy, Pandas, SciPy, Matplotlib, Scikit-learn, TensorFlow, PyTorch, PySpark), R, SQL, Tableau.

**Cloud Databases:** AWS S3, Sagemaker, EC2, GCP, Docker, MySQL, SQL Server, PostgreSQL

**ML Techniques:** Detection Classification Regression Clustering Algorithms, Ensemble methods, Hypothesis Testing, Statistical analysis

**Other Skills:** ETL Pipelining, Exploratory Data Analysis, Data Integration, Text Analysis, Git, A/B testing, Training, CI/CD, inference pipelines

**Soft Skills:** Collaboration, Presentation Skills, Problem-solving, Critical Thinking and Continuous Learning

## PROFESSIONAL EXPERIENCE

**Deloitte | Rosslyn, VA | Jan 2024 - Present**

**Data Analyst – Machine Learning R&D**

- Developed a graph based predictive anomaly detection system using Python on over 60,000 records, including **optimization, test & design of experiments and prescriptive analytics** to solve business problems.
- Designed **advanced ensemble framework using XGBoost, Random Forest and AdaBoost** to predict and validate over individual models and **enhanced prediction accuracy by 18%**
- Engineered down-sampling algorithm on high-frequency data in Python; 75% faster analysis and early detection of anomalies
- Created **advanced statistical models** using time series decomposition, hypothesis testing, PCA, and clustering to interpret anomaly trends and **dashboards in Tableau** that enhanced cross-team understanding and strategic insights for urban planners.
- Conducted **research on advanced machine learning methods to reduce false positives** in data anomalies, translating business requirements into technical solutions that resolved critical pain points
- Delivered **MLOps solutions using MLflow on AWS EC2** for automated deployment and monitoring, providing real-time anomaly streaming tracking to support data-driven decision-making

**Infosys | India | July 2020 – Dec 2022**

**Data Analyst**

- Facilitated collaboration between business users and technical team to streamline database performance, reducing error rates and improving data accuracy by 45% within cross-functional Agile teams using **Agile methodology**.
- Built **sales forecasting and demand planning models using ARIMA time series analysis**, increasing revenue by 20% product classification.
- Engineered ETL pipelines using **SQL queries and python scripts to preprocess large datasets from AWS S3**, improving data retrieval efficiency and utilizing **JIRA and Git** for version control and collaborative project management.
- Designed and executed A/B tests to evaluate feature performance and user engagement, analyzing results with statistical methods that contributed to a 15% increase in customer retention.
- Managed daily business-as-usual activities, including **database maintenance, dashboard troubleshooting, and cross-team collaboration** to ensure smooth workflow.

## ACADEMIC EXPERIENCE

- Optimized and **maintained ETL data pipelines using PySpark, integrating data from AWS S3 into Tableau dashboards** for real-time Climate predictions improved data flow reliability and enabled strategic financial decision-making.
- Built a **supervised MLOps pipeline to predict customer churn on a 10GB imbalanced CRM dataset** with 15K features, applying advanced preprocessing techniques with Pandas and NumPy, and improving recall and precision by 30%.
- Executed **data cleaning and preprocessing using Pandas and NumPy**, implementing advanced **missing data imputation techniques** to ensure high data quality and reliability for analytical models.
- Built an AI-powered customer support system with RAG architecture**, FAISS vector search, OpenAI API fallback, and a Streamlit chat UI, enabling real-time analytics and multi-format document uploads to enhance query accuracy and user engagement.
- Built a Streamlit AI app for **multimodal sentiment analysis using RoBERTa**, DistilRoBERTa for text, OpenAI CLIP for image sentiment, enabling **real-time text and image emotion detection and visualization with PyTorch and Plotly**.
- Designed and **deployed prediction pipeline in PyCharm using Python, Pandas, and SQL for data wrangling**, with scikit-learn models, robust feature engineering, and lineage tracking via MLflow and Git.
- Developed an **interactive data visualization dashboard using R, Plotly, and D3.js** transforming complex policy data into actionable insights that support targeted interventions and influence public safety strategies.

## EDUCATION

**Masters in Data Analytics and Engineering | George Mason University**

**Bachelor's in computer science and engineering | Vignana Bharathi University**