Srihas Gullapalli

gsrihas@iu.edu | (930) 333 1367 | linkedin.com/in/srihas-gullapalli | github.com/srihasgullapalli

Education

M.S. in Data Science Aug 2024 – May 2026
Indiana University Bloomington, Bloomington, IN
B.Tech. in Computer Science and Engineering **GPA: 8.7 / 10.0**
Sathyabama University, Chennai, India Jun 2020 – Jun 2024

Experience

Vosyn Inc. — Research Analyst Intern (Language Intelligence Project) Sep 2025 – Present
Remote / AI Communication Analytics

- Engineered multilingual data pipelines processing voice and text logs to improve VosynCore's cross-lingual communication models
- Developed NLP-based metrics to measure cultural and emotional adaptation, boosting translation model accuracy by 20%
- Applied clustering and semantic similarity algorithms to segment multilingual user behavior, enhancing personalization in voice/text delivery
- Partnered with AI researchers and product teams to operationalize analytics dashboards supporting real-time language adaptation insights

Projects

Event-Driven PII Redaction Pipeline — Spark & Kafka (Dockerized) Jul 2025 – Sep 2025

- Built a distributed ETL pipeline using Apache Kafka, Spark Streaming, and Docker, reducing end-to-end latency by 35%
- Automated regex + SHA-256 anonymization ensuring GDPR/CCPA compliance, cutting manual audits by 25%
- Designed monitoring modules for fault-tolerant ingestion, maintaining 99% uptime and improving data reliability

Advanced Credit Card Fraud Detection — XGBoost & Isolation Forest Feb 2025 – Mar 2025

- Developed ensemble ML models to improve fraud recall by 40% and reduce false positives by 18% using SMOTE and GridSearchCV
- Performed threshold calibration and SHAP analysis achieving AUC 0.96 and interpretability for stakeholders
- Automated preprocessing pipelines improving model reproducibility and deployment speed by 20%

Supply Chain Forecasting & Cost Optimization — Prophet & Power BI Sep 2024 – Nov 2024

- Built Prophet/ARIMA/XGBoost ensemble models to forecast demand with 27% higher accuracy
- Designed Power BI dashboards and SQL alerts that reduced logistics costs by 15% and improved planning efficiency
- Deployed forecasting automation on AWS EC2 reducing manual updates by 40%

Technical Skills

Programming: Python, R, SQL
Cloud: AWS, GCP, Azure
Frameworks & Libraries: Scikit-learn, TensorFlow, PyTorch, Pandas, NumPy
Data Engineering: Apache Kafka, Spark (Batch/Streaming), Docker, Git
Analytics & BI: Power BI, Tableau, Jupyter
Machine Learning: NLP, Supervised/Unsupervised ML, Time Series, Feature Engineering, Model Evaluation

Certifications

Infosys Springboard — Data Science with Python, Statistics for Data Science, Data Science Tools
NPTEL — Programming in Java