

# Sravan Kumar Uppoju

## Senior Data Scientist

+1 (636) 254-7650 | [sravankumaruppoju2@gmail.com](mailto:sravankumaruppoju2@gmail.com) | St Louis, MO | [LinkedIn](#) | [Github](#)

### PROFESSIONAL SUMMARY

**Senior Data Scientist** with 5+ years of experience designing and deploying production-ready Machine Learning and Generative AI solutions across Healthcare, Banking, Retail, and Service domains. Proficient in Python, SQL, R, Spark, and cloud platforms like AWS, Azure, GCP, with hands-on expertise in end-to-end ML lifecycle: data ingestion, feature engineering, model training, deployment, monitoring, and production support. Skilled in LLMs, Transformers, Retrieval-Augmented Generation (RAG), Prompt Engineering, Agentic AI, and MLOps pipelines. Experienced in time-series forecasting (ARIMA, Prophet, LSTM), recommendation systems, churn prediction, A/B testing, and secure API integration. Committed to Responsible AI using SHAP, AIF360, and Fairlearn for explainability, bias detection, and compliance.

### SKILLS

- Programming:** Python, SQL, R
- ML and AI:** Supervised Learning, Unsupervised Learning, Deep Learning, NLP, Feature Engineering, Model Evaluation, Time Series Forecasting, ARIMA, Prophet, LSTM, A/B Testing, Churn Prediction, Recommendation Systems, Predictive Modeling, Hyperparameter Tuning, Model Validation, Feature Selection
- Generative AI and Agentic AI:** GPT-4, GPT-3, BERT, LLaMA, Retrieval-Augmented Generation, Prompt Engineering, LangChain, FAISS, LangGraph, MCP, Multi-Agent Systems, Embeddings, Knowledge Graphs
- Frameworks and Libraries:** NumPy, Pandas, Scikit-learn, TensorFlow, PyTorch, Keras, Hugging Face, SciPy
- Big Data and Data Engineering:** Apache Spark, Hadoop, Hive, Kafka, Airflow, ETL Pipelines, Data Modeling, SQL Optimization, Feature Engineering, Data Ingestion, Streaming Pipelines, Data Warehousing
- Databases:** SQL Server, MySQL, Oracle, MongoDB, Snowflake, Neo4j
- Visualization and BI:** Power BI, Tableau, Advanced Excel, KPI dashboards
- Cloud:** AWS EC2, S3, Lambda, IAM, Azure Data Factory, Data Lake, Databricks, GCP Vertex AI, BigQuery
- MLOps and DevOps:** MLflow, Docker, CI/CD, Model Deployment, Data Ingestion, REST APIs, OAuth 2.0, Model Monitoring
- Methodologies:** Agile, SDLC, Git, GitHub, Jupyter Notebook
- Responsible AI and Compliance:** SHAP, AIF360, Fairlearn, HIPAA
- OS:** Windows, Linux

### EXPERIENCE

**Senior Data Scientist | Anheuser-Busch | NYC, NY** Aug 2024 – Present

- Built production-grade **ML, AI, and Generative AI pipelines**, performing RCA, defect triaging, redaction, and SLA monitoring in a Palantir Foundry-like environment.
- Developed **LLM-based RAG pipelines** using GPT-4, LangChain, and FAISS for enterprise document search and summarization bots.
- Implemented **Agentic AI workflows** with LangGraph and MCP, enabling memory-augmented, multi-agent conversational systems.
- Designed and maintained **time-series forecasting models** (ARIMA, Prophet, LSTM) on Azure Databricks with MLflow and Grafana, achieving **87% accuracy**.
- Built **churn prediction and recommendation models** (XGBoost, Random Forest), improving customer retention by **12%** and engagement metrics.
- Developed end-to-end **MLOps** pipelines across AWS, Azure, GCP (Vertex AI, BigQuery) with CI/CD automation using Jenkins and Tekton, reducing release cycles by 30%.
- Integrated **secure REST APIs** with **OAuth 2.0 authentication**, enabling controlled access to external AI services OpenAI, Hugging Face.
- Performed large-scale **data processing and feature engineering** using **PySpark, Spark SQL, Hive**, and streaming pipelines.
- Implemented **A/B Testing frameworks** for ranking and targeting models, improving response relevance by **23%**.
- Ensured **responsible AI compliance** by integrating **SHAP and AIF360** for explainability and bias detection.
- Developed **enterprise Power BI dashboards** with **DAX and RLS**, integrating Generative AI generated insights and reducing report refresh time by **40%**.
- Optimized **SQL queries** using indexes, CTEs, and execution plans, reducing runtimes from **90s to 8s**.
- Designed **graph-based AI solutions** using **Neo4j and knowledge graphs** for entity linking and semantic search.
- Collaborated with **product, marketing, and engineering teams** to translate business requirements into scalable AI solutions.

**Data Scientist | NantHealth | India** May 2022 – May 2023

- Designed **ETL and NLP pipelines** for healthcare claims, eligibility, and clinical data, handling **structured and unstructured datasets**.
- Developed **HIPAA compliant workflows**, redacting sensitive PII and integrating healthcare NLP models via **secure APIs** like FHIR, HL7, X12.
- Built **RAG and MCP pipelines** and **multi-agent AI workflows** for context-aware document retrieval and claims automation.

# Sravan Kumar Uppoju

## Senior Data Scientist

+1 (636) 254-7650 | [sravankumaruppoju2@gmail.com](mailto:sravankumaruppoju2@gmail.com) | St Louis, MO | [LinkedIn](#) | [Github](#)

- Applied **predictive modeling, statistical analysis, and ML and NLP techniques** to improve operational efficiency and accuracy.
- Developed **interactive Tableau dashboards** to visualize claims and operational KPIs, accelerating stakeholder decision-making.
- Explored **Generative AI models** Claude, LLaMA for document summarization and internal AI agents.
- Implemented **model explainability and fairness** using **SHAP and Fairlearn**.
- Collaborated cross-functionally to deliver analytics solutions that improved process efficiency and supported strategic business decisions.

**Data Engineer | Bank of America | India** **Oct 2019 – Mar 2022**

- Built and optimized **SQL-based ETL pipelines** and reporting workflows for multi-platform financial data.
- Developed Power BI and Tableau dashboards for internal teams, implementing RLS and usage tracking for adoption metrics
- Conducted **customer segmentation, CLV analysis, and profitability modeling**, helping optimize client strategies.
- Provided **L1 and L2 production support** and resolved data pipeline and BI report issues, improving SLA compliance.
- Automated KPI monitoring and reporting workflows to **reduce manual effort by 30%**.
- Collaborated with finance, marketing, and operations teams to **translate complex datasets into actionable insights**.
- Developed ad-hoc analytics and **trend analysis reports** for internal and client-facing stakeholders.

**Junior Data Analyst | Avon Technologies | India** **Apr 2019 – Oct 2019**

- Built **analytics pipelines** for clinical and product data using Python, Hive, Hadoop, and Spark SQL.
- Developed predictive models using **Python and R** to analyze product adoption, market trends, and operational efficiency.
- Designed dashboards in **Power BI and Tableau**, enabling real-time visualization of KPIs and business metrics.
- Processed **large-scale datasets 1TB or more**, performing aggregation, transformation, and cleaning for analysis.
- Automated data ingestion and preprocessing workflows, reducing manual effort and errors.
- Conducted **exploratory and statistical analysis** to identify trends and insights for marketing and product teams.

## EDUCATION

**Master of Science in Computer and Information Sciences**, Saint Louis University, St Louis, MO, May 2025

## CERTIFICATIONS

- [Microsoft Certified: Azure Data Engineer Associate](#)
- [IBM Data Science Professional Certificate](#)
- [Microsoft Certified: Power BI Data Analyst Associate](#)
- [AWS Certified Cloud Practitioner](#)