



# SIRICHANDANA BOMMA

+1-940-843-6795

[sc.bomma@gmail.com](mailto:sc.bomma@gmail.com)

[LinkedIn](#)

[GitHub](#)

[Portfolio](#)

## Professional Summary

Data Scientist with **3.5 years of experience** applying advanced analytics, machine learning, NLP, and generative AI to deliver business impact across healthcare, SaaS, and finance domains. Skilled in **Python programming and scripting**, building data pipelines, and applying **LLM-based solutions** to extract insights from structured and unstructured data. Experienced in deploying predictive models on **Azure ML, Databricks, and AWS SageMaker**, ensuring scalable and production-ready delivery. Strong background in **data visualization** with PowerBI, Tableau, and Streamlit to communicate insights and accelerate decision-making.

## Skills

**Programming & Scripting:** Python, Python Scripting, SQL, Bash, .NET, PostgreSQL, MySQL, MongoDB, BigQuery

**Data Engineering & Processing:** ETL/ELT Pipelines, PySpark, Dask, Airflow, Delta Lake, Data Modeling, Feature Engineering

**AI/ML & GenAI:** Machine Learning (Supervised & Unsupervised), NLP (spaCy, Hugging Face), LLMs (OpenAI APIs, LangChain, LangGraph), Generative AI Pipelines, MLflow, Predictive Modeling

**MLOps & Deployment:** MLflow, Docker, Kubernetes, FastAPI, Flask, REST APIs, AWS SageMaker, Azure ML, CI/CD Pipelines

**Visualization & BI Tools:** Tableau, Power BI, Streamlit, Plotly, Matplotlib, Seaborn — Interactive dashboards, KPI tracking, Executive reporting

**Cloud Platforms:** AWS, Azure, Databricks, Snowflake (familiar), GCP

**Automation & Optimization:** Workflow Monitoring, Data Validation, CI/CD, Python Scripting for Automation, DVC, Performance Tuning

## Experience

### Spok Inc

Jul 2024 – Present

Data Scientist

McKinney, TX

- Designed and deployed generative AI pipelines integrating **LLMs (OpenAI, Hugging Face)** with structured data, cutting query resolution time by 40%.
- Built end-to-end data pipelines in **Python, PySpark, and Airflow**, increasing ingestion throughput by 45% and lowering latency for healthcare analytics.
- Engineered **Python scripting solutions** for automation of ETL processes, reducing manual intervention by 30%.
- Developed and fine-tuned **NLP models** for clinical text processing, improving entity extraction and classification accuracy by 25%.
- Implemented Retrieval-Augmented Generation (**RAG**) pipelines and integrated **AI Agents (LangChain, LangGraph)** to automate compliance checks, reducing workload by 40%.
- Created **visualization dashboards** with PowerBI and Streamlit, enabling leadership to monitor KPIs and make data-driven decisions 30% faster.
- Deployed ML and GenAI models in **AWS SageMaker** with HIPAA compliance, ensuring secure and production-ready delivery.
- Authored technical documentation, conducted knowledge transfer, and mentored offshore teams, accelerating onboarding by 50%.

### Signify Health

May 2023 – Jun 2024

Data Scientist

Dallas, TX

- Built ML-integrated pipelines in **Python and Databricks** for patient risk scoring, improving predictive accuracy by 20% and reducing care gaps.
- Automated healthcare data workflows using **Python scripting and Airflow**, reducing reporting delays by 35%.
- Designed **NLP pipelines** for extracting medical entities from EHR text, achieving 92% accuracy in clinical data enrichment.

- Migrated legacy workflows to **Azure ML pipelines**, reducing deployment cycles from weeks to 3 days.
- Applied anomaly detection techniques in **PySpark and MLlib**, reducing false positives in analytics by 35%.
- Built preprocessing frameworks for large-scale EHR datasets, lowering reporting time by 30% and ensuring ML model readiness.
- Delivered **PowerBI dashboards** that visualized patient risk and compliance trends, cutting review time by 25%.
- Collaborated with clinical teams to transform raw data into compliance-ready insights, improving operational efficiency across healthcare programs.

#### Zensar Technologies Limited

Mar 2022 – Jul 2022

Junior Software Engineer

Hyderabad, India

- Developed **.NET utilities** to automate financial reporting, reducing manual workload by 25%.
- Enhanced legacy .NET applications with optimized data access layers, improving response times by 20%.
- Collaborated with QA and support teams to resolve production defects in .NET services, ensuring 99% uptime.

#### VITA Technologies

Aug 2021 – Feb 2022

Software Engineer Intern – Data Engineering & AI/ML

Warangal, India

- Designed proof-of-concept ETL pipelines with Python/SQL, improving client onboarding efficiency by 25%.
- Applied clustering and dimensionality reduction techniques for churn analysis, generating actionable retention insights.
- Built AWS Lambda serverless workflows for near real-time analytics, reducing reporting latency significantly.
- Developed **Streamlit** annotation and visualization tools, increasing labeling throughput by 4x and accelerating model training.

### Key Projects

#### AgentFabric – Autonomous Multi-Agent Research Assistant

Prototype: <https://agentfabric-etazidrjdbpg4ccc3mmqun.streamlit.app/>

- A collaborative AI framework where multiple specialized agents work together to automate research and analysis tasks, significantly reducing manual effort and accelerating insights.

#### RAGChain – Contextual Document Retrieval & QA Assistant

- An AI-powered system that combines retrieval-augmented generation with scalable pipelines to provide precise answers from large document sets, improving knowledge access and decision support.

#### MedAgent – Healthcare Compliance & Regulation Summarizer

Prototype: <https://medagent-3pudaugc7v2ayxychv6ecn.streamlit.app/>

- A healthcare compliance assistant that automatically processes and summarizes lengthy regulatory documents, helping professionals cut review time by more than half while ensuring accuracy.

### Education

#### Master of Science

University of North Texas

Information Science (2022-2024)

3.7 GPA

#### Bachelor of Technology

JNTU Hyderabad, India

Computer Science & Engineering (2017-2021)

3.4 GPA

### Certifications

- AI in Healthcare – Stanford University
- AI Agent Developer – Vanderbilt University
- Salesforce Certified AI Associate
- IBM Certified Machine Learning Developer
- Prompt Engineering – Vanderbilt University
- Certified Kubernetes Application Developer (CKAD)
- IBM Certified Enterprise Design Thinking Practitioner
- IBM Enterprise Design Thinking – Team Essentials for AI
- Data Science with Python – SimpliLearn
- Large Language Models – SimpliLearn
- Generative AI – SimpliLearn