

SRI RAM SATHIYA NARAYANAN

Boston, MA | Ph: +19497928260 | sathiyannarayanan.sr@northeastern.edu | [LinkedIn](#) | [Portfolio](#) | [Blog](#) | [GitHub](#)

SKILLS:

Languages: Python (EDA, ML, API development), R (statistical modeling), SQL (CRUD, query optimization)

Machine Learning & AI: Apache Spark MLlib, BioBERT, prompt engineering, retrieval-augmented generation (RAG)

Big Data & Cloud: Apache Spark, Databricks, Zeppelin notebooks, AWS (EC2, Lambda, S3), Docker, GitHub Actions, CI/CD

Data: MySQL, MongoDB, Power BI, Excel

Frameworks, Libraries & Backend Tools: Flask, Node.js, LangChain, HuggingFace Transformers, FastAPI, RESTful API, Postman

Tools & Collaboration: Git version control, JIRA, Confluence, SAP ERP

Conceptual Skills: Data preprocessing, data governance, statistical analysis, semantic search, model evaluation

Certifications: Azure Fundamentals, Azure AI Fundamentals, IBM Specialization in AI & IoT, IBM Enterprise Design Thinking, IBM AI Analyst, AWS Certified AI Practitioner

Education:

- Master of Science in Computer Science**

May 2027 (Expected)

Khoury College of Computer Sciences, Northeastern University, Boston, MA.

Coursework: Algorithms, Programming Design Paradigm, Database Management Systems, Web Development, Machine Learning, MLOps, Foundations for Generative AI

- Bachelor of Technology in Computer Science and Engineering**

June 2022

B.S. Abdur Rahman Crescent Institute of Science and Technology, India.

Coursework: Statistical Methods for Data Analysis, Software Engineering, Graph Theory, Data Visualization, Artificial Intelligence and Machine Learning, Web Analytics and Social Media Mining, Software Project Management

WORK EXPERIENCE:

Consultant - Ernst & Young, Chennai, India.

October 2024 - December 2024

- Led requirement-gathering and process mapping sessions with supply chain stakeholders, converting insights into technical requirements that reduced process bottlenecks by **20%**, ensuring smooth analytics solution delivery.

Associate Consultant - Ernst & Young, Chennai, India.

October 2023 - September 2024

- Developed and deployed advanced demand forecasting models using **R, Python, and Power BI**, improving procurement planning and contract stability, leading to a measurable **20%** boost in procurement efficiency.
- Drove inventory optimization by analyzing stock turnover in Power BI, SAP MM and R with advanced statistical techniques, reducing wastage by **10%** and boosting availability above **97%**, minimizing losses and enhancing supply resilience.

Senior Analyst - Ernst & Young, Chennai, India.

July 2022 - September 2023

- Designed pricing analytics solutions using Python and Excel, integrating historical sales and market trends to enable data-driven decision-making, negotiations, reduce manual effort, and accelerate procurement cycles by **20%**.
- Automated data extraction, transformation, and reporting workflows, reducing manual preparation time by **30%** and delivering actionable insights for pricing, forecasting, and procurement optimization.

Intern - Ernst & Young, Chennai, India.

February 2022 - July 2022

- Collaborated with IT, Finance, and Procurement to gather requirements, analyze business processes, and deliver functional specifications supporting customer-focused scalable digital transformation initiatives.
-

PROJECTS:

- Biomedical RAG Summarization Pipeline** (Python, LangChain, HuggingFace, FAISS, BioBERT)

October 2025

Implemented a Healthcare specific retrieval-augmented generation (RAG) pipeline to process biomedical documents and generate multi-level summaries (patient, abstract, clinical) reducing manual effort by 40%, included intelligent document chunking using LangChain's TextSplitter and section-aware tagging for context-preserving processing. Applied semantic embeddings with BioBERT and integrated FAISS vector database for high-accuracy similarity search and retrieval.

- FMCG Lakehouse ETL Pipeline** (Databricks, Python, SQL, Apache Spark, AWS S3, Power BI)

October 2025

Developed an end-to-end Databricks ETL pipeline to integrate datasets after a major retail acquisition. Implemented ingestion workflows from Amazon S3 into a Medallion architecture using Spark, Python, and SQL. Derived unified Gold-layer tables supporting BI dashboards with consolidated sales, product, and operational analytics.

- Text-to-SQL Chatbot** (Python, LangChain, Multi-LLM, FastAPI, MySQL, Excel)

September 2025

Designed an end-to-end Text-to-SQL chatbot allowing natural language queries on MySQL databases, preprocessed Excel datasets for ingestion, created a multi-LLM LangChain pipeline for SQL generation, validation, and optimization, and deployed a FastAPI backend to handle requests and return structured, real-time responses.

- Big Data Regression using Spark & Zeppelin** (Apache Spark MLlib, Zeppelin, Databricks)

August 2025

Built end-to-end regression pipeline using Spark MLlib; performed data ingestion, feature engineering (StringIndexer, VectorAssembler), and model evaluation. Automated cloud job scheduling and caching with Zeppelin; achieved 93% model accuracy using Spark SQL and DataFrames.

- Regional Sales Analytics Dashboard** (Python, Power BI, Pandas, NumPy, Excel)

June 2025

Created BI pipeline for multi-source sales data, conducted EDA to identify seasonal and regional trends. Built interactive Power BI dashboards driving 30% sales growth via SKU optimization and targeted promotions.