


Rajesh Subramanian

Senior Big Data | AI | ML Engineer | Full-Stack Software Architect

✉ srajesh2712@gmail.com ☎ +44 7887 176510 📍 Stirling, United Kingdom

🌐 <https://github.com/srajesh2712>  <https://www.linkedin.com/in/srajesh27/>

🎧 <https://www.kaggle.com/rajeshsubramanian27>  <https://srajesh2712.github.io>



Profile

- Specialized in designing **high-throughput distributed systems** with **metadata-driven ETL frameworks** reducing pipeline onboarding time by **70%**
- **Results Driven** Architectural Strategy: Proven track record of transitioning complex monolithic systems to **Cloud-Native Microservices** (Python / Java / Spring Boot / Go ecosystem) with **12 Factor App** methodologies and **Software Design Principles** increasing operational efficiency **gains by 27%**
- **Cost Efficient** Distributed Data Engineering: Architected large-scale pipelines using Big Data Technologies such as **Hadoop, Spark, Docker / Kubernetes (GKE)**, and **Airflow** etc. optimizing workloads to achieve **40% reductions in compute costs**
- **Business Efficiency** with AI & Emerging Tech: Improved business performance by 20% by implementing Machine Learning & Neural Network models
- **Leadership & Full-Stack Mastery**: Deep proficiency in Python and Java ecosystems (in **Spring Boot, FastAPI, Go**) , Proven track record of ability to upgrade along with evolution of technology
- Active **opensource contributor** in Kaggle and github projects
- **Communication**: B1+ Level German skills.

Skills

Big Data & Analytics

Apache Spark, Apache Iceberg, Parquet, Hadoop, Kafka, Airflow, Sedona, Dask, Lakehouse

Programming & Backend

Python, Java J2EE, Spring Boot, Microservices Go, Gin, Fast API, Node, Express JS, Nest JS

Front End

React JS, Angular

AI | Generative AI | ML

Machine Learning, Neural Networks, Model Context Protocol (MCP), Building Agent, Retrieval-Augmented Generation (RAG), **HNSW (Vector Search Indexing)**, LangChain, Natural Language Processing (NLP), Computer Vision

Databases & Graph

PostgreSQL, MongoDB, Neo4j, ArangoDB, Oracle, MySQL, Teradata, Database Migration

Cloud & DevOps

Cloud-Native Architecture (AWS, Google Cloud), Docker, Kubernetes, Infrastructure as Code (Terraform), Circle CI

APIs & Messaging

Kafka, Kafka Streams, Google Cloud Pub/Sub, Active MQ, gRPC, REST, Webhooks, Socket Programming

Work Experience

Lead Data Engineer

07/2025 – Present | Chennai, India

BigThinkCode Technologies

- Architecting projects with techstack React, NestJS, Postgres
- Designing Pipelines for scalable distribution systems
- Socket Programming for LIVE PCM Audio transfer for podcast

Data & AI Intern

11/2025 – 12/2025 | Stirling, Scotland

Keep Scotland Beautiful

- Participated in the architecture design phase of the AI project, providing input on data scalability and near-real-time processing requirements
- Actively contributed to the Litter Detection Project by collecting and labeling a high-quality dataset of **1000** images for AI model training

Lead Fullstack / Data Engineer

08/2022 – 01/2025 | Chennai, India

BigThinkCode Technologies 

- **Architected a Metadata-Driven ETL Framework:** Developed a pluggable "Configuration-over-Code" system using **PySpark**, reducing new pipeline onboarding time by **70%**
- **Engineered Kubernetes-Native Infrastructure:** Deployed Spark on **GKE**, implementing dynamic dependency loading for distributed archives (ZIP/JAR) to ensure seamless scaling and a **40% reduction in compute costs**
- **Decoupled Business Logic from Engine:** Established a 3-tier architecture (JSON Manifest, Orchestrator, Worker Logic) that increased code reusability and maintained **99.9% pipeline reliability** across global brand datasets
- **Architected** scalable Enterprise level application using **SOLID** Principles and **12 Factor App** using Fast API / React / Postgres / AWS

Full Stack Engineer

03/2021 – 07/2022 | Pondicherry, India

Software - Freelancer 

Technical Manager

03/2012 – 03/2021 | Pondicherry, India

eSales Technologies Pvt Ltd

- **Architected Scalable Cloud-Native Web Applications** by Leveraging **Java / Python ecosystem** to design a microservices-based e-commerce ecosystem, improving system modularity and reducing feature deployment cycles by **40%**
- **Engineered Generic E-commerce Data Warehousing:** Built a modular data warehouse architecture to centralize cross-platform retail data, utilizing **metadata-driven** patterns to support rapid business intelligence and reporting.
- **Modernized Legacy Infrastructure:** Spearheaded the migration of monolithic PLB legacy applications to a **Cloud-Native architecture**, resulting in a **27% gain in operational efficiency** and significantly lower maintenance overhead.
- **Optimized High-Performance APIs:** Redesigned core system interfaces using **FastAPI** and **Gin**, achieving a **30% reduction in latency** and ensuring the platform could scale to meet peak traffic demands.

Associate Consultant

12/2010 – 03/2012 | Chennai, India

Oracle Financial Software Services

Developed and implemented core banking modules for Tier-1 financial clients, ensuring compliance with regulatory standards.

Programmer Analyst

10/2009 – 12/2010 | Chennai, India

*Cognizant Technology Solutions India Pvt Limited***Senior Software Engineer**

11/2006 – 09/2009 | Chennai, India

iGate Global Solutions

Led the successful migration of two legacy systems onto a unified platform in terradata, improving data consistency and reporting speed by **27%**

Education

M.Sc Big Data

01/2025 – 09/2026 | Stirling, UK

*University of Stirling***M.B.A - International Business**

08/2008 – 05/2012 | Pondicherry, India

Loyola College/ Pondicherry University [↗](#)**Bachelor of Technology - Electrical & Electronics Engineering**

08/2002 – 05/2006 | Pondicherry, India

Pondicherry University

Certifications

- AWS Cloud Practitioner [↗](#)
- Astronomer Certification DAG
Authoring for Apache Airflow [↗](#)
- Azure AI Fundamentals