

Raj Shekhar Singh

Bangalore, India • +91-8840082361 — rajsingh170901@gmail.com — GitHub — LinkedIn

EDUCATION

PES University

Bachelor of Technology (B.Tech) in Electrical and Electronics Engineering

2021 – 2025

Bangalore, Karnataka

EXPERIENCE

Infosys Limited

Specialist Programmer Level 1

Jun 2025 – Sep 2025

Mysuru, India

- Architected and enhanced ETL pipelines on Azure Data Lake Storage Gen2 using Apache Spark and Kafka, achieving 35 percent lower latency and 40 percent higher throughput.
- Streamlined Spark- and Kafka-based real-time data processing, cutting latency by 1/3rd while boosting throughput.
- Configured scalable data workflows integrating MongoDB NoSQL Database, MySQL Relational Database, and other NoSQL systems, reducing query latency by 40 percent.
- Implemented Generative Artificial Intelligence models for analytics automation, cutting reporting time by 30 percent.
- Collaborated with Agile Scrum teams with Continuous Integration / Continuous Deployment (CI/CD) pipelines, accelerating deployment cycles by 25 percent.

Surat Construction Pvt. Ltd.

Jan 2025 – Jun 2025

Data Analytics Intern

Lucknow, India

- Owned and optimized SQL-driven transactional databases for construction assets, raw materials, procurement, and project cost records, improving analytical query performance by up to 35 percent.
- Developed Python-based analytics and visualization pipelines using Pandas and Matplotlib to deliver actionable insights on inventory utilization, asset efficiency, and project progress for management teams.

PROJECTS

AI-Powered Portfolio Website | React.js, Spring Boot, MongoDB, Redis, Groq AI, Resend

[Live Demo](#)

- Full-stack portfolio with AI chat assistant; enforced Redis caching strategy reducing response times by 95% and enabling sub-millisecond portfolio data retrieval with intelligent TTL-based cache invalidation.
- Designed 25+ REST APIs with JWT auth, MongoDB document storage, async email system (Resend), and WebSocket real-time chat; deployed on Railway & Netlify with CI/CD pipelines achieving 99.5% uptime and 100+ concurrent users.

Customer 360 Analytics Platform | MongoDB, Apache Spark, Kafka, Azure Data Lake, Hadoop HDFS

[Live Demo](#)

- Unified e-commerce, mobile, and IoT data from 500K+ users; orchestrated real-time Kafka-Spark pipelines reducing ingestion latency by 45% and processing 1TB+ daily via Azure Data Lake and Hadoop HDFS.
- Engineered data warehouse with Hive supporting scalable analytics and churn prediction models achieving 85% accuracy, boosting user retention by 20%.

AI Driven Analytics Platform | FastAPI, LangChain, PostgreSQL, FAISS, BM25, Groq, Next.js

[Live Demo](#)

- Built a production-grade AI analytics platform allowing users to upload raw CSV datasets and obtain instant business insights through natural-language conversations, eliminating manual SQL and dashboarding workflows.
- Engineered a hybrid RAG backend using FastAPI, BM25, and FAISS with PostgreSQL-backed multi-tenant storage and JWT security, delivering low-latency, context-aware analytics with interactive visualizations on a Next.js frontend.

Task Manager Dashboard | Spring Boot, Django, Java, JavaScript, Bootstrap, REST API, AI Suggestions

[Live Demo](#)

- Deployed a full-stack task management application with CRUD functionality, task priority and status tracking, image handling via the Pixabay API, and an AI-powered suggestion chatbot using Hugging Face on Render.com.

TECHNICAL SKILLS

Programming Languages: Java, Python, C, SQL, JavaScript, HTML, CSS, Shell Scripting

Backend Frameworks: Spring Boot, Spring Security, Spring MVC, Django, FastAPI, REST APIs, Microservices

AI and LLMs: LangChain, FAISS, Groq AI, GPT-4, RAG, Vector Embeddings, TensorFlow, Generative AI, Agentic AI

Frontend: React.js, Framer Motion, TailwindCSS, Bootstrap

Databases: MongoDB, MySQL, ChromaDB, Vector Databases, NoSQL, Query Optimization, Data Modeling

Caching and Performance: Redis, In-Memory Optimization, Index Management

Big Data: Apache Spark, Kafka, Hadoop HDFS, Hive, PySpark, Azure Data Lake, ETL Pipelines

DevOps and Cloud: Docker, Railway, Netlify, Git, CI/CD, Monitoring, WebSocket

LEADERSHIP & CERTIFICATIONS

Operations Team Lead – Organized 60+ teams with 25% outreach growth — Samarpana'23 Disciplinary Committee — Coordinated 1500+ participants at state-level marathon — Top 30 – CyberSecurity Hackathon — Certifications: IBM Python for Data Science (Aug 2023), British Airways Data Science Virtual Experience (Jul 2023), HackerRank Competitive Programming Certification (Intermediate, 2023).