

SHIVAM GALVE

Software Engineer | Kurukshetra, Haryana, India

☎ +91 9416818306

✉ shivamgalve5@gmail.com

🌐 [linkedin.com/in/shivamgalve](https://www.linkedin.com/in/shivamgalve)

📁 Portfolio: dezter.com

Experience

Xperium.ai

Jan 2023 – Present

Software Development Engineer

Gurugram, India

- Delivered full-stack features and backend services for retail clients including Titan, owning development through deployment and production support across scalable **distributed systems**.
- Designed and developed **microservices** in Python and TypeScript with well-defined service boundaries; wrote **unit and integration tests** for all services, ensuring reliability and regression coverage across deployments.
- Architected and optimized **RESTful APIs** for high-throughput backend services with **load balancing** and horizontal scaling, reducing average response time by **40%** and improving system throughput by **3x**.
- Scaled marketing campaign processing to **10,000 campaigns/day** and **1M+ email, WhatsApp, and SMS deliveries**; implemented retry logic and event-driven workflows, reducing delivery failures by **35%**.
- Reduced Kubernetes cluster RAM usage by **50%** by analyzing service behavior, separating responsibilities, and optimizing long-running processes; enabled **high availability** and zero-downtime deployments.
- Migrated core services from GCP to AWS, improving deployment reliability by **30%** and reducing infrastructure downtime during releases by **90%**.
- Built an in-product AI chatbot powered by an MCP server, leveraging **RAG** with embeddings-based retrieval and reranking for analytics Q&A and guided dashboard navigation over structured business data.
- Optimized **Redis-based caching** and background workloads, reducing p95 latency by **45%** and eliminating resource contention spikes under peak load.
- Resolved production issues including race conditions, stuck states, and incorrect routing, improving system stability and reducing incident rate by **60%**.

Project

Graph-Based Cognitive Retrieval System

Next.js, TypeScript, FastAPI, LangChain, LlamaIndex, PostgreSQL, Redis, FAISS, Neo4j, AWS, Docker

- Architected a graph-based cognitive retrieval platform combining semantic vector search, knowledge graphs, and contextual memory to deliver adaptive **RAG** across structured and unstructured data.
- Engineered hybrid retrieval pipelines using embeddings, graph traversal, temporal context, and personalized memory for context-aware querying, recommendations, and conversational intelligence.
- Developed a real-time streaming AI interface with contextual citations, persistent memory, and scalable backend infrastructure using FastAPI, Redis, PostgreSQL, FAISS, and Neo4j deployed via Docker on AWS.

Technical Skills

Languages & Frameworks: JavaScript, TypeScript, Python, Next.js, React, Node.js, FastAPI, Flask

Databases & Messaging: PostgreSQL, MongoDB, Redis, Neo4j, Pub/Sub, Event-Driven Architecture

AI/ML & Vector Search: RAG, Embeddings, LangChain, LlamaIndex, FAISS, MCP

Systems & Architecture: Distributed Systems, Microservices, Load Balancing, REST APIs, Caching, Unit Testing

Cloud, DevOps & Tools: AWS, GCP, Docker, Kubernetes (K8s), Linux, Git, GitHub

Education

JC Bose University of Science and Technology, Faridabad

August 2019 – June 2023

Bachelor of Technology (CGPA of 8.18)

Haryana, India

Achievements

- Mastered **Data Structures and Algorithms** by solving **800+ problems** across LeetCode, CodeChef, and Codeforces, with strong command over graphs, dynamic programming, trees, and system-level patterns.
- Diagnosed and resolved critical production incidents including race conditions, stuck states, and routing failures, directly improving system uptime and client SLAs.
- Self-taught distributed systems and cloud-native architecture; applied concepts like **Redis-based caching**, pub/sub, and Kubernetes optimization directly in production at scale.

Coding Profiles

LeetCode: leetcode.com/shivamgalve5

Codeforces: codeforces.com/profile/shivamgalve

CodeChef: codechef.com/users/shivam.galve

GitHub: github.com/shivamgalav