Harsh Kothari

Software Engineer with 3+ years building large-scale distributed systems and cloud services
Phone: 8722140114 — Email: harshkothari12312@gmail.com

Education

University of Illinois at Chicago (UIC), Chicago, IL *Expected May 2026*
Master of Science in Computer Science

Manipal Institute of Technology, Manipal, India *2017–2022*
Bachelor of Technology in Computer & Communication Engineering

Skills

Languages: C#, Java, C++, Python, JavaScript/TypeScript, Scala, SQL
Cloud & Distributed Systems: AWS (EC2, S3, RDS, EMR, EKS, Lambda), distributed architecture, microservices
Backend Development: .NET Framework, Spring Boot, Node.js, RESTful APIs, design patterns
DevOps: CI/CD (Jenkins, GitHub Actions), Docker, Kubernetes, monitoring, observability
Databases: PostgreSQL, MySQL, MongoDB, Redis, Neo4j, database optimization
Data Processing: Apache Spark, Apache Flink, Apache Kafka, distributed data pipelines

Work Experience

Software Engineer *Sep 2021 – May 2024 (3 years)*
E-Tech Innovative Services Private Limited, Jaipur

- Led end-to-end development of E-Krishi Mandi marketplace, architecting and deploying large-scale distributed system serving **100,000+ users** and processing millions of transactions with cloud-based infrastructure ensuring high availability and reliability.
- Designed and implemented scalable microservices architecture using .NET Framework, Java, and Node.js with RESTful APIs, applying design patterns to ensure maintainability, performance, and scalability across distributed services.
- Built production DevOps pipelines with automated CI/CD (Jenkins), comprehensive testing strategies, and deployment automation, reducing deployment time by **40%** while maintaining **99.9%** uptime.
- Optimized application performance by **30%** through systematic profiling, implementing caching strategies (Redis), database optimization, and monitoring systems to improve observability and efficiency.
- Implemented monitoring and observability solutions for distributed services, debugging production issues, and ensuring consistent operations at scale.
- Conducted code reviews, mentored junior engineers, and established coding standards that reduced production bugs by **85%**.
- Collaborated with product managers and stakeholders to translate business requirements into technical specifications, designing APIs and delivering scalable solutions that improved operational efficiency by 30%.
- Integrated third-party services (payment gateways, SMS APIs) with robust error handling, retry mechanisms, and monitoring for production reliability across distributed systems.

Graduate Research Assistant *Feb 2025 – Present*
University of Illinois Chicago

- Optimized machine-unlearning pipeline by **20%** through systematic refactoring of activation paths and implementing efficient verification algorithms in Python and C++, reducing end-to-end proof and inference time on CNN benchmarks.
- Parallelized machine unlearning using SISA (sharding/slicing) to retrain only impacted components, improving deletion efficiency and scalability through thread-level parallel execution for shard retraining.
- Implemented deletion-impact mapping table: for each training point, computed the set of dependent samples whose predictions change after removal, enabling targeted unlearning analysis and reproducible deletion workflows.
- Built cluster-intersection table over flip-sets to identify shared vulnerable samples, characterizing instability patterns and enabling systematic analysis of model behavior across data clusters.
- Built reproducible experiment harnesses with deterministic seeds, consistent data splits, automated logging, and evaluation pipelines to fairly compare unlearning methods and track performance metrics across experimental configurations.

Projects

GraphRAG: Distributed Knowledge Graph Pipeline for LLMs *Sep 2025 – Dec 2025*
Scala, Apache Spark, Apache Flink, Neo4j, Delta Lake, AWS (EMR, EKS, S3), Akka HTTP, Kubernetes

- Designed and implemented large-scale distributed system processing **500+ research papers**, architecting two-phase pipeline with incremental delta-indexer (Spark + Delta Lake) and real-time processing (Flink + Neo4j), reducing reprocessing overhead by **70%**.
- Built scalable cloud-based REST API microservices with 8 endpoints for semantic search and graph traversal, implementing observability, error handling, and monitoring achieving sub-2-second query latency over 4,300+ relationships.
- Implemented distributed stream processing with Apache Flink for real-time data transformation and entity extraction, handling event-driven workflows with exactly-once semantics and fault-tolerant state management.
- Deployed cloud infrastructure on AWS EKS with Kubernetes orchestration, Docker containerization, CI/CD automation, and monitoring dashboards—demonstrating full software lifecycle from design to production deployment and operations.
- Applied software engineering best practices including code reviews, comprehensive testing, versioned deployments with blue-green strategies, ensuring high availability, reliability, and performance at scale.