Rishabh Sood

Linkedin: linkedin.com/in/rishabhsood2/
Github: github.com/RishabhSood

Email: sood.rishabh@hotmail.com
Mobile: +91-8920492376

EDUCATION

Georgia Institute of Technology — GPA: 4.0
MS Computer Science, Machine Learning — *2025 - Present*
Courses: *Deep Learning, Machine Learning, Robotics: AI Techniques, Cognitive Science*

Thapar Institute of Engineering & Technology, Patiala — GPA: 9.92
B.E. Computer Science & Engineering — *July 2019 - June 2023*

EXPERIENCE

Apple — Hyderabad, India
Software Engineer (Full Time) — *Jan 2024 - Present*

- **Support App Chat (Generative Experience)**: Engineered one of Apple's first generative multi-turn conversational experiences for customer support, impacting millions of users with an instant, 24/7 channel.

  -Implemented a multi-agentic architecture with tool-calling capabilities to perform complex actions like checking warranty status, scheduling repairs, and querying knowledge base articles.
  - Engineered a robust Retrieval-Augmented Generation (RAG) pipeline to ground responses in Apple's proprietary knowledge base, dramatically reducing hallucinations and increasing response groundedness.
  - Implemented safety guardrails and content filters to ensure generated responses aligned with Apple's brand and quality standards.

- **Apple Developer Search**: Led the geographical expansion to multiple regions and languages.

  - Overcame language barriers to deliver a consistent user experience on the developer website.
  - Optimized storage costs by 90% (to 1/10) without compromising search quality.

- **Knowledge Hub**: Architected and engineered a centralized data extraction-transformation-storage pipeline to unify data operations, featuring a custom plugin architecture for dynamic pipeline structuring (Sources: Message Queues/APIs; Transformation: Cleaning, Chunking, Embedding; Storage: Vector/Relational DBs).

- **Tech-Stack & Skills**: RAG, GraphRAG, Python, GRPC, LLM (Train / FineTune / Deploy), Textgrad, MCP, Knowledge Graphs, FastAPI, Multilingual Search, Apache Solr, OpenSearch, Neo4J, Docker

Goldman Sachs — Hyderabad, India
Analyst (Full Time) — *Jul 2023 - Dec 2023*

- **Internal App Store**: Contributed to the overhaul of central App Store by migrating the entire architecture from Windows servers to Kubernetes and introducing a regression test suite to automate UAT.
- **Data Aggregator:** : Led the development of a versatile data aggregator, architecting a highly configurable system to automate Kubernetes job scheduling for data lake onboarding and significantly reducing costs.

- **Tech-Stack & Skills**: Java, Spring Boot, Kubernetes, Docker, Gitlab CI-CD Pipelines, Gherkin, Cypress

Apple — Hyderabad, India
Software Development Engineer (Intern) — *Jan 2023 - Jun 2023*

- **Enhanced Apple Support Chat Experience:** : Developed an Intelligent Recommendation Service using the dynamic Apple Support Knowledge Base (real-time sync), reducing response time from minutes to a fraction of a second.

  - Improved user support by integrating generative responses and sentiment analysis.
  - Innovated a centralized package to reduce code redundancy and enhance productivity.

- **Tech-Stack & Skills**: Python, PyPI Packaging, FastAPI, MongoDB, OpenSearch, Solr, Kubernetes, Docker, Generative AI, Amazon EKS.

Goldman Sachs — Bengaluru, India
Summer Analyst (Intern) — *Jun 2022 - Jul 2022*

- **Emergency Call Notification Service:**: Developed & deployed a notification service on a global scale.
  - Replaced the legacy 3rd party vendor solution, thereby **saving USD 35-40K / annum**.

  - **Successfully delivered** emergency call notifications to the global security team **in real time**.
  - The implemented solution **easily extends to new regions** by modifying the configuration file.

- **Tech-Stack & Skills**: Golang, Kubernetes, Prometheus (Metrics & Alerting), Grafana, Fluentbit, BigQuery, Docker, Gitlab CI-CD Pipelines, SRE.

TRAININGS & CERTIFICATIONS

Generative AI With Large Language Models, AWS Fundamentals Specialization, Architecting on AWS, Amazon EKS, Tensorflow in Practice Specialization, Deep Learning Specialization, Google Code-In

SKILLS SUMMARY

- **Languages:**   C/C++, Java, Python, Rust, Golang, HTML, CSS, JS
- **Frameworks:**   Langchain, RIG(Rust), Spring Boot, FastAPI, GRPC, Poem(Rust), ReactJS, PyTorch, Django
- **DevOPS:**   Kubernetes (AWS EKS), Docker, Spinnaker
- **Database:**   SQLite, PostgreSQL, Redis, MongoDB, OpenSearch, QDrant, Apache Solr, Neo4J