

RUSHIL KUMAR

+91-9999346778 | rushilkumar01@gmail.com | rushilkumar.com | [LinkedIn](#) | [GitHub](#) | Bengaluru, India

EXPERIENCE

Software Engineer | Cisco Systems

Jul 2024 – Present

Mercury: Kubernetes-based observability platform adopted by 2 major ISPs and 4 test teams. Registers device clusters, ingests logs via Fluentd/Loki, deduplicates and stores metrics via Thanos/Prometheus, and triggers alert-driven mitigation via Alertmanager.

- Engineered a multi-agent AI system for root-cause analysis that placed top 10/~100 at Cisco's Innovation Hackathon, then rolled it into Mercury, killing the manual investigation loop by auto-correlating evidence across scattered data
- Owned AutoResync, automating the recovery steps a system once needed after a failure, validated at 128,000-subscriber scale and saving 10+ minutes of manual fix-up work per incident
- Architected a scheduled monitoring microservice that runs diagnostic checks so engineers can triage issues and monitor system health while testing new features, killing the one-off scripts users once wrote and maintained themselves
- Launched IPv6 support to meet a major Indian ISP requirement (validated at 1M-subscriber scale) and hardened the service with config hot-reload, storage health metrics, and a full circuit breaker lifecycle

Technical Intern | Cisco Systems

Jan 2024 – Jun 2024

- Automated subscriber redistribution across hardware partitions in Python, enforcing an 8k-subscriber capacity limit across 16,000 subscribers; created a React Native monitoring dashboard for 3 network providers, cutting operator investigation time 30%

DevOps Intern | Reeudo (30-member, Seed-Funded Cross-Border Payments Startup)

Jun 2023 – Jul 2023

- Automated live FX rate ingestion (Google Apps Script, stored in AWS S3), ensuring accurate real-time exchange rates for 1,000+ users on a platform processing cross-border transactions; secured dashboard data for 400+ consultants using client-side AES encryption (CryptoJS) on AWS Lambda and wired up Jenkins CI/CD for Flutter deploys to a 1,000+ customer base

SDE Intern | MapmyIndia

Jun 2022 – Jul 2022

- Slashed multi-modal routing API error rate by 50% (Valhalla engine, 5,000+ users) and improved JSON parsing efficiency by 40% via a C++ nlohmann wrapper, eliminating repeated deserialization overhead on Transitland REST API calls

PROJECTS

NewsFlow AI | newsflow.in | Python · OpenRouter · Google Cloud Neural2 TTS

Apr 2026 – Jun 2026

- Built a configurable multi-agent pipeline that ingests any email newsletter and delivers a personalized daily podcast via Google Cloud Neural2 TTS, cutting repeat stories 35% via semantic dedup
- Removed the single point of failure in model access by chaining Anthropic, multiple OpenRouter models, and local Ollama, auto-blacklisting any model that fails or hits its limit so no provider outage breaks the pipeline
- Isolated each stage behind a typed contract and checkpoint so failures retry alone, keeping the pipeline failure free for 60+ days

Pinchit | AI-agent workflow platform | pinchit.in | Go · LLM Orchestration

Feb 2025 – Jul 2026

- Owned LLM-based semantic task routing for Pinchit's multi-agent dev platform, killed the old keyword-overlap matcher, and launched a 4-stage autonomous pitchdeck pipeline that ran live in 2 client demos and closed 1 deal
- Built a 13-tool agentic GitHub integration (PR/issue/branch lifecycle): spec, plan, code, tests. Processed ~15 PRs/issues in prod
- Forked an open-source agent-agency framework and fine-tuned 4 bundles of 5 agents each across engineering, marketing, UX, and research, covering the business's full range of operational needs

TruthSeek | AI product-research startup | truthseek.in | Python · Playwright · Supabase

Aug 2025 – Jan 2026

- Scaled interview throughput 100x, running 100 calls in parallel in under 30 minutes versus a PM's 10 hours for 20 calls manually
- Shipped human-in-the-loop live-call monitoring, letting a researcher watch or jump in and take over an AI interview mid-call
- Established calling-agent guardrails on wording, mentions, intent exposure, and follow-up depth to keep interviews compliant
- Integrated a product intelligence dashboard: scrapes 500+ reviews/product across 10 ASINs daily via BrightData proxies, then feeds a map-reduce summarization pipeline with hallucination-blocking validation for FMCG growth teams
- Designed penalty-adjusted scoring (5x weight on 1-star reviews) so rankings reflect satisfaction instead of rating × volume

TECHNICAL SKILLS

- Languages:** Python, Go, C/C++, JavaScript/TypeScript, SQL, Dart
- Backend & Infra:** FastAPI, NestJS, Flask, Docker, Kubernetes, PostgreSQL, AWS (S3, Lambda), Jenkins CI/CD, Linux
- AI/ML:** Agent orchestration, multi-agent systems, voice-agent workflows, guardrails, LLM reliability, vector search, ABSA
- Observability:** Prometheus, Thanos, Grafana, Alertmanager, Fluentd, Loki, MongoDB, SQLite, Supabase
- Frontend & Mobile:** React Native, Flutter, Firebase

EDUCATION

- Thapar Institute of Engineering and Technology, B.E. Electrical Engineering, CGPA 8.11/10
- DPS Vasant Kunj, New Delhi, XII Science (CBSE), 92.4%

Jun 2024
2020