

RISHAB DUGAR

+91-9330549405 | rishabdugar.tech@gmail.com

linkedin.com/in/rishabdugar | github.com/kingrishabdugar | rishabdugar.is-a.dev | leetcode.com/rishabdugar

SUMMARY

Software Engineer with ~3 years at Accenture building production backend microservices, event-driven data pipelines, and LLM-powered agentic AI systems (GenAI) on AWS and Azure. Triple cloud-certified (AWS DVA-C02, GCP Professional Machine Learning Engineer, Azure AI-102). On-call DRI for client-facing services at 99.9% uptime; promoted within 2.5 years.

EXPERIENCE

Accenture

Bengaluru / Kolkata, India

Data & AI Senior Analyst — Backend & AI Platform Engineering (promoted from Analyst, Dec 2025)

Aug 2023 – Present

- Built and operated production **Python and Node.js microservices** on AWS Lambda + API Gateway with Cognito/JWT auth, serving **2,000+ daily enterprise users** at **99.9% uptime** and **sub-500ms p95 latency**.
- Re-architected a **15+ microservice platform** into a modular serverless design with clean REST contracts and shared handler layers, cutting infrastructure cost **~70%** and release cycle from weeks to minutes.
- Designed an **event-driven analytics pipeline** (WebSockets, Kinesis Firehose, Parquet, Athena) with dual-write durability and per-file fault tolerance — **10x compression**, **~200x lower query cost** across 19 analytical views.
- Delivered production conversational AI (chat + voice) for US state healthcare clients handling **~150K calls/month**, with multi-layer LLM guardrails (off-topic and language detection, response filtering) — **15K+** self-service interactions and **~35% adoption lift** post-launch.
- Engineered an **omnichannel AI voice agent** (Amazon Nova Sonic, Amazon Connect, Bedrock Claude) unifying IVR, WebRTC, and chat with DynamoDB session continuity — cut per-call LLM inference cost **~70%** at sub-2s p95 voice latency.
- Cut voice authentication latency **~60%** via parallel DynamoDB lookups (**ThreadPoolExecutor**), LRU caching, bounded retries, and circuit breakers.
- Served as **on-call DRI**: built centralized logging on OpenSearch + Kibana, authored automated runbooks, and drove incident triage and RCAs under client SLAs — **MTTR reduced ~60%**.
- Shipped a RAG-based test-case generator (~96% accuracy) cutting QA effort ~90%; maintained Docker + ECR + App Runner CI/CD pipelines; drove code reviews and RCAs; mentored 30+ engineering interns and co-authored architecture decision records.

PROJECTS

Universal Agentic Platform | *FastAPI, React, MCP, LangGraph, DynamoDB, AWS App Runner*

Internal deployment

- Built a multi-tenant agent platform (20+ agents, 40+ MCP tools) sustaining **~1,000 concurrent users** with SSE streaming and a 4-layer memory design (in-memory, SQLite checkpoints, DynamoDB, file sessions); tool-change detection rebuilds agents at runtime without losing conversation history.

AI-Enabled Invoice & Commerce Platform | *FastAPI (async), React, SQLite WAL, WebAuthn, Linux*

[Live](#)

- Engineered an async backend on a **2 GB single-worker Linux host**: tiered store (SQLite WAL source of truth + JSON mirror) with automatic failover and **ordered write-replay on recovery**; semaphore-serialized 150–300 MB PDF renders to prevent OOM.
- Added WebAuthn biometric auth (public-key only, sign-count clone detection), enumeration-safe capability-token e-sign flow with constant-time verification, and health-gated CI/CD via a self-hosted GitHub Actions runner.

SLM Fine-Tuning Pipeline | *Unsloth, PEFT-LoRA, 4-bit Meta LLaMA 3, Google Colab*

[Medium](#)

- Fine-tuned LLaMA 3 with 4-bit PEFT-LoRA quantization on free-tier Colab; published a step-by-step guide with evaluation metrics and benchmarks.

TECHNICAL SKILLS

Languages: Python, Java, JavaScript/TypeScript, SQL, C++

Concepts: Data Structures & Algorithms, System Design, Distributed Systems, Concurrency & Multithreading, Scalability, Caching

Backend: FastAPI, Node.js, REST API design, microservices, event-driven architecture, async I/O, WebSockets, OOP, design patterns

GenAI & Agents: LLMs, Generative AI, Agentic AI systems, AWS Bedrock, LangChain, LangGraph, MCP, RAG (hybrid search + re-ranking), prompt caching, Pinecone, ChromaDB

Cloud & Infra: AWS (Lambda, API Gateway, DynamoDB, Kinesis, Athena, Cognito, S3), Azure Functions, Docker, CI/CD, Linux, Git

Data & Caching: DynamoDB, PostgreSQL, MySQL, Redis, OpenSearch, Parquet

Reliability & Observability: circuit breakers, retries, fault tolerance, CloudWatch, OpenSearch/Kibana, distributed tracing, unit & integration testing, incident response

EDUCATION

National Institute of Technology, Durgapur

CGPA: 8.72 / 10.0

B.Tech (Bachelor of Technology), Electrical Engineering

Aug 2019 – Jun 2023

CERTIFICATIONS & ACHIEVEMENTS

Certifications: [AWS Certified Developer – Associate](#) | [Google Cloud Professional ML Engineer](#) | [Azure AI Engineer Associate \(AI-102\)](#) | [Accenture Data Architect Associate](#)

Competitive Programming: [LeetCode Knight](#) — 1983, top 2.84% (450+ solved, 50 contests); CodeChef 4-star (max 1892); TCS CodeVita S10 — Global Rank 1,979 / 136,000+; Amazon ML Summer School 2022 — top 0.5% of 19,000+

Recognition: Accenture Multi-Skilled Champion FY25; 2x iChamp Client Focus Award; Top 1% LinkedIn Programming Voice FY 2024–25 (13k+ followers)