

DARPAN PATEL

904-962-3760 | dpcodesintech@gmail.com | linkedin.com/in/darpan-patel-094b151a1 | github.com/dpcodestech

SUMMARY

Software Engineer with 3+ years building scalable, distributed backend systems and AI/LLM-integrated applications for high-volume financial platforms. Deep expertise in Java, Spring Boot, and microservices; cloud experience spans Azure in production and a self-proposed, self-led AWS migration (EC2, Elastic Load Balancing, Auto Scaling) that cut server response time 80%, plus ongoing AWS certification study. Applied AI expertise includes LLM integration, prompt engineering, and model fine-tuning. Proven track record shipping systems that process 10M+ daily records with sub-second latency for Bank of America via Infosys.

TECHNICAL SKILLS

Languages: Java (Expert), Python (Expert), Dart (Flutter), SQL

Backend & Cloud: Spring Boot, Microservices, REST/GraphQL APIs, Apache Kafka, AWS (EC2, Elastic Load Balancing, Auto Scaling, Lambda, S3), Azure, GCP (Cloud Scheduler), Docker, Kubernetes, CI/CD (GitHub Actions, Jenkins)

AI & ML: LLM Integration (GPT-4/4o, Gemini, HuggingFace), Prompt Engineering, Model Fine-Tuning, BLEU/ROUGE Evaluation, Gradio

Databases & Tools: MySQL, Firestore, NoSQL Design, Git, Jira, Postman

PROFESSIONAL EXPERIENCE

Technology Analyst — Infosys (Client: Bank of America) | Newark, DE

Feb 2026 – Present

- Built and scaled a Java/Spring Boot microservices platform for real-time financial data ingestion, architecting the pipeline to process 10M+ records/day at sub-second latency in support of core transaction-reporting systems.
- Designed and deployed a real-time event-streaming layer using Apache Kafka to replace legacy batch syncing, delivering high availability and seamless data consistency across globally distributed environments.
- Built and rolled out Azure-based deployment automation and monitoring tooling, cutting manual release effort and lifting backend resource efficiency 20% for ML-driven analytics workloads.
- Redesigned a legacy flow that accepted all input types (~5,000 requests/day, only ~1,000 of which were the valid type) to accept only the required input — eliminated unnecessary request volume and grew valid-request throughput to ~6,000/day; confirmed by the security team as improving system security posture (~20% per internal security standards review).
- Built an internal AI-powered search tool on top of the GitHub Copilot API that crawls and indexes internal documentation and web pages, replacing manual multi-page documentation searches with a single AI-generated answer — saved an estimated 5 developer-hours/day for early users; currently in pilot testing ahead of broader rollout.
- Stepped into an informal lead role for 4 weeks when the feature lead went on cooldown leave, coordinating development, business requirement-gathering, and QA support across a team reduced from 6 to 3 onshore engineers (contractors on a 3-month cooldown) — kept the August release and November development cycle on track.

Skills: Java, Spring Boot, Apache Kafka, Azure, Microservices, GitHub Copilot API, Cross-Functional Leadership

iOS Developer, Flutter — MicroFone (Startup, Part-Time Support)

Jan 2026 – Present

- Built an automated news-ingestion pipeline using a GCP Cloud Scheduler job that triggers the Brave Search API every 24 hours and writes results to Firebase, replacing a manual, direct-API call from the app — improved automation coverage 90% vs. the prior implementation.
- Introduced a branch-based deployment workflow (dev branch build → stakeholder review/sign-off → merge to main) for the team, replacing ad hoc direct-to-production pushes.
- Redesigned and rebuilt the app's live navigation menu, replacing a cache-based implementation — improved menu response time 40%.
- Re-architected image-heavy submission pages to pull from Firebase into a local cache with on-demand backend refresh instead of re-fetching every load, improving page performance 70%.
- Contributed core features for the app's App Store launch.

Skills: Flutter, Dart, GCP Cloud Scheduler, Firebase, Brave Search API, iOS, Git branching workflows

Software Engineer — Infosys (Client: Bank of America) | Richardson, TX Dec 2022 – Aug 2024

- Built distributed backend services in Java that processed 10M+ financial records daily, powering downstream ML analytics pipelines.
- Re-engineered the microservice concurrency model with low-latency threading, lifting system throughput 35%.
- Built an automated schema-validation and preprocessing layer for ML data pipelines, eliminating malformed-data failures and raising reliability for downstream consumers.
- Built and rolled out a suite of 150+ automated unit/integration tests, cutting production defects 15%.

Skills: Java, Microservices, System Design, Concurrency, CI/CD, Unit/Integration Testing

Software Engineer — Revature | Reston, VA

Jun 2022 – Dec 2022

- Proposed and led migration of an internal project's on-premises server to AWS (EC2, Elastic Load Balancer, Auto Scaling) as an alternative to purchasing additional on-prem hardware — distributed load across US regions and improved server response time by 80%.
- Built CI/CD automation pipelines for backend workflows, accelerating AI feature rollout by 25%.
- Designed and built contract-validation modules for ML components, ensuring schema integrity across microservices.
- Diagnosed and resolved 20+ critical production bugs, stabilizing AI-integrated modules.

Skills: AWS (EC2, Elastic Load Balancer, Auto Scaling), CI/CD, Contract Validation, Java

PROJECTS

| Novora — AI Companion App (LAHacks 2025)

- Built and shipped a real-time AI companion app on Gemini Pro and multimodal APIs, using intelligent token batching to cut inference cost while keeping responses fast.
- Designed sub-second conversational flows with custom caching and streaming responses, eliminating perceptible lag in live demo conditions.
- Built the cloud infrastructure layer — Firebase Auth plus modular AI state providers — to support scalable multi-user sessions.

| AI-Based Slogan Generator — GPT-2 Fine-Tuned Model

- Fine-tuned GPT-2 on a curated slogan dataset, achieving 69% BLEU — outperforming baseline models on the same benchmark.
- Built a custom evaluation pipeline (BLEU, ROUGE, perplexity) with visual dashboards to track model quality across iterations.
- Built and deployed an interactive Gradio web UI enabling real-time experimentation and inference tuning.

EDUCATION

California State University, Los Angeles — MS in Computer Science (2024–2025)

Navrachana University — BTech in Information Technology (2018–2022)