Aviral Garg

Vancouver, British Columbia, Canada ✉ aviral.garg@icloud.com in/aviralgarg 🌐 aviralgarg.com/

SKILLS

- **Core**: System Design, Python, Flask, FastAPI, Node.js, React, AWS (CDK, CloudFormation, CloudWatch, Step Functions, Lambda, S3, EC2, X-Ray, ECS, DynamoDB, Athena, Glue, SNS, SQS, SageMaker, API Gateway), Azure, GCP, Docker, CI/CD.
- **AI**: openAI-agents-python, Ollama, Agentic Orchestration, AI agents, Agents Framework, MCP Servers, RAG, Synthetic Data
- **ML**: EDA, Feature Engineering, NLP, Vector Embeddings, Vector Databases, LLM, Time-Series, tsfresh, TPOT, Prophet
- **Integrations:** OpenAI, Anthropic, Gemini, Deepgram, Yahoo Finance | **Tools:** Cursor, Co-pilot, JetBrains

EXPERIENCE

Software Development Engineer
Amazon - Beauty Tech Team   February 2021 – March 2025, Vancouver, BC

- Led a cross-functional team of 8 engineers and data scientists to deploy AWS Step Functions serverless data pipelines with 50% reduced execution time. Pipelines automatically scaled for growth processing 40+ TB of product catalog data daily and powering Makeup Virtual Try-On experience for 40M+ monthly customers across 15+ countries.
- Built internal-documentation-powered AI agents to provide instant answers to CI/CD admin questions, identify documentation gaps, and flag outdated content. Improved question resolution time by 82%, with projected savings of 240+ dev-hrs/month from integration with the internal admin ticketing system.
- Built 4-staged CI/CD python build-chains with granular system monitoring, and bottleneck tracing using CloudWatch & X-ray. Received org-wide accolades for robust documentation and operational readiness for every step of the SDLC.
- Led an organization-wide documentation workshop, establishing a systematic platform for subject-matter-expert reference and reaching 100% completion of product and CI/CD documentation of the Beauty Tech Organization.
- Spearheaded CI/CD-driven store-page-builder project for Amazon clients using React and Python.
- Integrated external microservices and A/B testing framework increasing client customization capabilities by 300%.
- Automated monthly business reporting for Beauty Recommendations Team reducing manual reporting time by 90%.
- Achieved 100% stakeholder satisfaction across all reporting consumers through automated system implementation.
- Won first place at Amazon 2022 Hackathon with Augmented Reality iOS makeup-tutorial app leading a four-person team.

Software Development Engineer
Amazon - Prime Pantry Team   June 2020 – February 2021, Vancouver, BC

- Architected robust Kotlin applications achieving 100% test coverage for Prime Pantry API development.
- Served in weekly on-call rotation to troubleshoot complex build, release and developer-tooling issues that blocked the team.
- Cut setup time by 75% by automating developer onboarding and environment setup.
- Saved 3,528 developer hours annually across a 7-person team while eliminating setup-related errors.
- Optimized API performance, reducing page-load impact by 37% and increasing add-to-cart conversion rates by 2.5% that directly contributed to over 3 million monthly revenue for the Prime Pantry platform.
- Championed AWS CDK adoption across development teams for standardized CI/CD pipelines.
- Built automated CI/CD pipelines, saving 12+ developer-hours weekly and cutting deployment errors by 95%.
- Delivered $30K annual cost savings through systematic AWS infrastructure optimization and usage policy implementation.
- Led organization-wide CodeGuru Profiler adoption across 7 projects reducing on-call incidents and maintenance hours.

AI Software Engineer
T4G Limited   October 2018 – December 2019, Vancouver, BC

- Delivered enterprise-grade chatbots with full CI/CD, leveraging Azure DevOps, using Node.js, Cosmos DB, and SQL.
- Optimized promo-code generation achieving 40% efficiency improvement enabling 4 additional high-value contracts.
- Deployed, managed and tested physical Windows-computers for Impark/HangTag parking enforcement vehicles.
- Automated extraction of 10 years of git version history from 15 applications saving 7 dev hours weekly with Python scripts.
- Architected a Named-Entity-Recognition chatbot proof-of-concept using NLP, directly securing a $100K contract.
- Demonstrated AI capabilities to enterprise clients establishing company's reputation in the AI/ML space.
- Directly contributed to doubling the department's yearly profits, driving company revenue growth, and market expansion.
- Developed mission-critical C# desktop applications and SSIS data integration packages serving 50+ million users globally.

Software Engineer

Texavie Technologies Inc.   **December 2017 - May 2018, Vancouver, BC**
- Engineered cross-platform mobile, desktop, and mixed-reality headset proof-of-concepts in Unity3D.
- Demonstrated wearable prototypes securing investor interest and validating product-market fit for AR/VR technologies.
- Developed critical prototype debugging tools reducing calibration time by 60% enabling on-time product launch.
- Saved weeks of development effort while ensuring quality standards for consumer-ready AR/VR devices.
- Architected cross-platform game development pipelines for Windows, macOS, and Linux doubling target customer base.

PROJECTS (2025)

Bob
- Built a multi-threaded personal voice assistant with 8 operational modes for hands-free development and accessibility.
- Reduced Agentic AI platform costs by 40x through extensive prompt engineering and optimization techniques.
- Implemented zero speech data loss during processing with real-time speech-to-tool-calls conversion.
- Created an accessibility solution for developers with back pain, visual impairments, or typing difficulties.
- Integrated OpenAI Whisper for local ML-based speech recognition with real-time audio processing pipeline.

Live Transcriber (7.8K downloads)
pypi.org/project/livetranscriber
- Deployed a multi-threaded open-source Python package with real-time speech transcription and zero data loss.
- Architected a single-file WebSocket wrapper around Deepgram API with async/await supporting sync and async callbacks.
- Designed a flexible configuration system that allowed overriding Deepgram parameters while maintaining Nova-3 defaults.
- Implemented a comprehensive resource management with automatic cleanup, graceful shutdown, and error handling.
- Created a seamless PyPI distribution with proper dependency management enabling pip installation and integration.

Pre-bunker health communications system
github.com/gaviral/pre_bunker_health_communications_system
- Built AI agents for health misinformation prevention, simulating public reactions, generating evidence-backed prebunks.
- Created a comprehensive prototype with implementation across 19 versions achieving 65-80% risk reduction.
- Built a 5-stage pipeline: Claim Extraction, Risk Assessment, Audience Simulation, Evidence Validation, Countermeasures.
- Developed 12 specialized agent personas for comprehensive audience simulation with multi-source evidence validation.
- Created a FastAPI web interface with async processing and real-time risk assessment.
- Developed a novel Misinterpretability@k metric for quantifying health communication risk.

Resume Coach
coach.aviralgarg.com
- Built full-stack GPT-powered AI app with LangChain on AWS Lambda, API Gateway, DynamoDB, S3, and CloudFront.
- Created React 19 + TypeScript frontend with CI/CD pipeline using GitHub Actions, deploying production application.
- Implemented session persistence with DynamoDB TTL and contextual follow-up chat system.

FastAPI Stock Gainers
- Built a real-time stock market tracking application with FastAPI and SQLAlchemy ORM supporting PostgreSQL & SQLite.
- Integrated Yahoo Finance API for real-time stock market data with comprehensive user authentication system.
- Created an admin dashboard for database management with responsive Bootstrap UI and color-coded visualization.
- Deployed the application in production on AWS with auto-refresh mechanism and custom stock symbol tracking.

EDUCATION

Machine Learning Switch Up Program
Interview Kickstart · 2025
- Exploratory Data Analysis, Supervised, Unsupervised & Reinforcement Learning, Neural Networks, Agentic Frameworks

B.A.Sc., Computer Engineering (Software Engineering Major)
University of British Columbia · Vancouver, BC · 2019