

Svadrut Kukunooru

kukunoorusvadrut@gmail.com | linkedin.com/in/svadrut | github.com/svadrutk

SUMMARY

Software Engineer building production ML systems at Wayfair. Reduced support costs by \$1M annually through LLM-powered automation serving 1K+ daily active users. AWS Certified Developer with hands-on experience in Python, distributed systems, and production ML deployment.

PROFESSIONAL EXPERIENCE

Wayfair

Boston, MA

Software Engineer, Knowledge Management, Legal Tech

June 2024 – Present

- Engineered LLM-powered support automation system processing 73K annual tickets with 70% deflection rate, directly saving **\$1M in operational costs**
- Built distributed session management infrastructure handling 1000+ concurrent users with 99.9% uptime using Redis caching, Firestore persistence, and custom retry logic
- Developed Python-based evaluation framework measuring LLM accuracy across 5 key metrics (hallucination rate < 5%, precision > 85%), enabling data-driven model selection
- Optimized BigQuery analytics pipeline processing 10K daily interactions, reducing query time by 60% through partitioning and materialized views
- Won company-wide innovation challenge (1st/40+ teams) by building Waybor: multimodal AI system converting CAD files to assembly instructions using GPT-4 Vision API

PROJECTS

TrackHuntr | Python, React

August 2025 – September 2025

- Engineered AI-powered DJ tracklist extraction system using Groq LLM (gpt-oss-20b) with 4-stage parallel pipeline, achieving 90%+ accuracy in identifying tracks from YouTube videos
- Built Chrome extension with real-time video synchronization and Spotify/SoundCloud integration, with seamless YouTube SPA navigation handling and React 19-based UI
- Implemented intelligent validation system with confidence scoring and duration verification to ensure extracted tracks match video timestamps

Campusfy | Python, Next.js

Feb 2024 – May 2024

- Built full-stack web application serving 5K+ students across 2 universities for intelligent course discovery. 10K+ impressions
- Implemented RAG system using Gemini embeddings and Pinecone vector database to search 20K+ course listings with < 200ms latency
- Developed custom ranking algorithm incorporating difficulty scores, professor ratings, and user engagement metrics

TECHNICAL SKILLS

Languages: Python, JavaScript/TypeScript, Java, SQL

ML/AI: PyTorch, Scikit-learn, LangChain, OpenAI API, Embedding models, Vector databases

Backend: FastAPI, Flask, Spring Boot, REST APIs, GraphQL

Cloud/Data: AWS (Certified Developer), GCP, BigQuery, PostgreSQL, Redis, Docker, Kubernetes

Testing/Tools: pytest, Jest, Git, CI/CD (GitHub Actions), Terraform, Datadog

EDUCATION

University of Wisconsin – Madison

Madison, WI

Bachelor of Science in Computer Science, Minor in Data Science

2021 – 2024