

SHUBHANG VAGVALA

Boston, Massachusetts | (413) 362-0915 | shubhang680@gmail.com

LinkedIn: linkedin.com/in/sv68 | **GitHub:** github.com/shubhangv68 | **Portfolio:** shubhangv68.github.io/pf.github.io/

SUMMARY

Software Engineer with an M.S. in Computer Science and experience building scalable backend systems, cloud services, and GenAI applications. Skilled in performance optimization, distributed systems, microservices, and full-stack development. Open to impactful roles in backend, full-stack, or GenAI engineering.

WORK AUTHORIZATION

US Citizenship - No Sponsorship Required

EDUCATION

University of Massachusetts Amherst

Feb 2024 – May 2025

Master of Science in Computer Science, GPA: 3.97

Coursework: Machine Learning, Artificial Intelligence, Distributed Operating Systems, Algorithms in Data Science

EXPERIENCE

MediBuddy

Jun 2022 – Dec 2023

Software Engineer

Bengaluru, India

- Led development and scaling of backend microservices using **Spring Boot** and **Java**, improving API throughput by **60%** and reducing order failures by **12%**.
- Built **asynchronous Redis-based APIs**, optimizing concurrency handling and reducing server response time and CPU load by **30%**.
- Spearheaded an AI-driven FAQ assistant with **OpenAI LLMs** and prompt engineering, cutting support resolution time by **35%** and boosting customer satisfaction.
- Delivered A/B-tested features using **React.js/Redux**, improving engagement with real-time promotional content on the user dashboard.
- Integrated **Google Maps APIs** for live logistics tracking, enhancing route accuracy and improving delivery efficiency by **15–20%**.

Swiggy

Jun 2021 – May 2022

Software Engineer

Bengaluru, India

- Developed pricing microservices in **Go**, **Kafka**, and **gRPC**, reducing latency by **25%** and increasing pricing consistency by **20%**.
- Designed resilient data pipelines with **Kafka consumers** and feature toggles, maintaining a **99.5% SLA** across business-critical services.
- Automated inventory audits via **Databricks** and **Apache Hive**, eliminating **45%** of ETL inconsistencies and minimizing manual errors.
- Built internal **Grafana** dashboards and observability tooling to reduce on-call triage time by **10%** through proactive alerting and traceability.

SELECTED PROJECTS

AI-Powered Research Paper Q&A Assistant

Feb 2025 - May 2025

- Designed a document Q&A assistant using **LangChain**, **FAISS**, and **OpenAI APIs**, featuring a **RAG** pipeline for chunking, semantic search, and prompt engineering; deployed via **FastAPI** and **Streamlit** to serve academic queries in real time.

BERT Powered Product Recommendation Engine

May 2024 - Jul 2024

- Implemented an NLP-based recommendation system using **PyTorch**, **BERT**, and **HuggingFace Transformers**; analyzed **10M+** product reviews, boosted cold-start performance, and showcased pipeline deployment techniques.

Real Time Scheduling & Communication Platform

Feb 2022 - May 2022

- Developed a real-time scheduling platform using the **MERN** stack, **Google APIs**, **SendGrid**, and **Redis**; supported **50+** concurrent sessions with consistently low latency (**<500ms**) and integrated scalable multi-service architecture.

E-Commerce Microservices Architecture

March 2021 - May 2021

- Crafted and containerized microservices using **Spring Boot**, **React**, and **MongoDB**; demonstrated scalable service communication, RESTful design, and modular deployment.

TECHNICAL SKILLS

Languages & Frameworks: Java, Go, Python, JavaScript (ES6+), SQL, Spring Boot, Node.js, React, gRPC, REST APIs, Express.js

GenAI & ML: OpenAI, HuggingFace Transformers, LangChain, BERT, PyTorch, Scikit-learn, FAISS

Cloud & DevOps: AWS, Azure, Docker, Kubernetes, Databricks, CI/CD, Jenkins, GitHub Actions

Data Engineering: Kafka, Redis, MongoDB, PostgreSQL, Apache Hive, ETL Pipelines

Other: Agile/Scrum, TDD, Microservices, System Design, Prompt Engineering, Observability (Grafana)