Harish Kumar Samala

Connecticut, USA — +1 (475)-208-0329 — samalaharishkumar99@gmail.com — linkedin.com/in/samala-harish-kumar — Portfolio

Professional Summary

Software Engineer focused on building web applications, RESTful APIs, and AI-powered solutions. Strong experience in Python and SQL for automating complex data processing and optimizing database performance. Hands-on experience with Generative AI using OpenAI APIs and Mistral LLMs for chatbot development, embeddings, and semantic caching. Experienced in Flask-based API development, CI/CD automation, data engineering, and full-stack application delivery. Collaborative team player with experience in code reviews, documentation, and cross-functional development.

Skills

Languages & Frameworks: Python, Flask, Django, Streamlit, LangChain, SQLAlchemy, Dash

LLMs & GenAI: Mistral, Groq API, OpenAI APIs, LLaMA 3, HuggingFace Embeddings, Conversational RAG, Semantic Search

Vector Stores & Retrieval: FAISS, Chroma, RecursiveCharacterTextSplitter

Databases: PostgreSQL, MySQL, MongoDB, SQL

Data Processing: Pandas, NumPy, PySpark

Web Scraping & Automation: Selenium, BeautifulSoup

Document Processing: PyPDFLoader

Cloud & DevOps: AWS, Azure, GCP, Jenkins, Azure Pipelines, Git, Azure DevOps, GitHub

Infrastructure & Testing: Nginx, JMeter, dotenv

Analytics & ML: Sentiment Analysis, Apriori Algorithm, Data Visualization

Work History

Advanced Knowledge Technologies
Python Developer — Jul 2024 – Present

- Developed and deployed a production-ready Conversational RAG system using LangChain, Groq LLMs, OpenAI and HuggingFace embeddings, and Streamlit, enabling contextual PDF-based chat with persistent memory and multi-turn reasoning and reducing support response times by 40%.
- Designed and optimized modular pipelines for document ingestion, vector embedding, and semantic search using FAISS and Chroma, improving retrieval precision by 35% and reducing latency by 25%.
- Implemented history-aware prompts and session continuity using LangChain tools and Mistral LLMs, enhancing multi-turn reasoning accuracy by 30%.
- Built robust text-cleaning pipelines with OpenAI APIs and enhanced PDF onboarding using PyPDFLoader to handle complex documents including forms and tables.
- Spearheaded development of a domain-specific generative AI chatbot that automated high-complexity workflows and saved approximately 10 hours per week in manual operations.
- Developed RESTful APIs using Flask and Django, integrated CI/CD pipelines with Azure DevOps, and optimized database performance through indexing and caching, reducing query execution time by 40%.
- Engineered automated web scraping and data transformation pipelines using Selenium, BeautifulSoup, SQL, Pandas, and NumPy to enable real-time analytics and business insights.

- Applied the Apriori algorithm to usage logs to extract user behavior patterns and improve recommendation relevance.
- Automated stakeholder notifications using Python smtplib, implemented error-handling pipelines, and contributed to Agile delivery through code reviews and documentation.
- Built internal dashboards using Python Dash and performed sentiment analysis on customer reviews to support data-driven product and service improvements.

Projects

SparkleClean – Full-Stack Service Booking Platform — 2025
- Developed a full-stack booking platform with customer and admin portals.
- Built a responsive frontend using React, TypeScript, Tailwind CSS, and React Router.
- Implemented secure admin dashboard with full CRUD operations and role-based authentication.
- Designed multi-step booking workflows and real-time booking status tracking.
- Built backend services using Java 17 and Spring Boot with RESTful APIs.
- Implemented booking number generation and booking status workflows.
- Used Spring Data JPA with MySQL for data persistence.
- Deployed backend on Railway and frontend on Vercel.

Superior Limousine LLC – Corporate Website — 2025
- Designed and developed a fully responsive website for a luxury transportation company.
- Built frontend using TypeScript and Vite.
- Implemented mobile-first layouts and accessibility-focused user interfaces.
- Integrated Framer Motion animations and optimized media delivery using Cloudinary.
- Applied SEO best practices and structured metadata.
- Deployed the site on Vercel and configured DNS.

Case Notes (Hospital) – NJIT
- Developed a system to summarize and analyze patient case documents.
- Built backend services using NestJS, TypeScript, Prisma, PlanetScale, and AWS Lambda.
- Designed APIs and tested integrations using Postman.

Evaluating Student Performance – NJIT
- Built machine learning models to predict student performance from examination data.
- Generated performance trend insights to identify strengths and improvement areas.
- Used Pandas, NumPy, and Matplotlib for data analysis and visualization.

Crop Yield Prediction – NJIT
- Built linear regression models to forecast agricultural crop yields.
- Analyzed weather, temperature, and pesticide datasets to improve prediction accuracy.
- Achieved a mean squared error of approximately 1.75.

Vulnerability Scanning Using Nessus – NJIT
- Performed network vulnerability scanning using Tenable Nessus.
- Generated reports and remediation recommendations to improve organizational security posture.

Certifications

- Data Structures and Algorithms with Python
- AWS Certified Cloud Practitioner
- Generative AI Foundations

Education

New Jersey Institute of Technology, Newark, NJ — May 2024
Master of Computer Science, GPA: 3.1/4.0

Siddhartha Institute of Technology (SITS), Hyderabad, India — Aug 2021
Bachelor of Engineering in Computer Science, CGPA: 6.9/10