SAI SOWMITH KANTIPUDI
Python Developer / Gen AI Engineer
Hyderabad, India | 9441563209 | sowmith.kantipudi@outlook.com | www.linkedin.com/in/sowmith-k

PROFESSIONAL SUMMARY

Python Developer and GenAI Engineer with 3 years of experience building scalable, data-driven applications. Expert in integrating LLMs using LangChain, OpenAI, and Hugging Face with modern Python web frameworks like Django and FastAPI. Proven ability to build full-stack solutions with React, design ML pipelines, and deploy NLP and generative models to production. Passionate about combining Python software engineering with cutting-edge AI/ML tools to deliver intelligent applications.

WORK EXPERIENCE

Software Engineer **January 2024 – December 2024**
Technodrive Business Solutions, Michigan, USA

- Developed enterprise-grade full-stack applications using **Python, Django, React, and PostgreSQL**, handling authentication, routing, and RESTful service layers.
- Implemented reusable backend modules in Django and **FastAPI** for user management, ticket processing, and audit logging.
- Designed and optimized PostgreSQL schemas and ORM models to support transactional operations and dynamic queries.
- Integrated **LLMs using OpenAI API and LangChain** to enable intelligent ticket classification and resolution suggestions.
- Architected and built vector search mechanisms with **FAISS and SentenceTransformers** to enhance retrieval-augmented generation workflows.
- Refactored synchronous endpoints into async workflows to reduce latency in real-time GenAI response systems.
- Developed custom Python APIs to manage orchestration between frontend input, LLM response generation, and database updates.
- Deployed containerized applications using **Azure Kubernetes Service (AKS)** and orchestrated CI/CD workflows through **Azure Pipelines** for scalable and automated release management.
- Collaborated with cross-functional teams to ship scalable AI-powered features and enhance platform responsiveness.

Application Development Analyst **August 2020 - June 2022**
Accenture, Hyderabad, India

- Enhanced scalable backend services using **Python, FastAPI, and Django**, supporting insurance policy management, claims routing, and user access flows.
- Built modular frontend components with React and Bootstrap, integrated with backend APIs to display dynamic policy and claim data.
- Implemented advanced PostgreSQL features such as materialized views, triggers, and stored procedures to support real-time reporting, event logging, and data integrity across insurance claim workflows.
- Created **RESTful APIs** for web clients and integrated external services using token-based authentication.
- Automated operational processes like claims ingestion and batch updates using Python scripts and cron jobs.
- Built ML microservices using Scikit-learn and FastAPI to predict fraud and classify incoming claims.
- Incorporated ML insights into dashboards, helping underwriters and agents prioritize high-risk policies.
- Oversaw end-to-end deployment on **AWS Fargate** and **ECS**, with integrated logging and alerting via **AWS CloudWatch** and infrastructure automation using **CloudFormation**.
- Used Docker for consistent development environments and GitHub Actions to automate testing, builds, and deployments.

TECHNICAL SKILLS

Programming Languages: Python, JavaScript, SQL, HTML5, CSS3
GenAI & NLP: LangChain, LlamaIndex, OpenAI API, Hugging Face Transformers, SentenceTransformers, RAG, Prompt Engineering, ChromaDB, FAISS
Machine Learning: Scikit-learn, TensorFlow, PyTorch, Pandas, NumPy, Matplotlib, Seaborn
Web & API Development: Django, FastAPI, Flask, Django REST Framework, JWT Auth

Frontend: React.js, Bootstrap,HTML,CSS
Data & Streaming: Apache Kafka, Spark Streaming, PostgreSQL, MySQL, MongoDB
Cloud & DevOps: AWS (SageMaker, EC2), Azure ML, Docker, GitHub Actions, Jenkins, Nginx
Visualization & Interfaces: Streamlit, Gradio
Other Tools: Git, Jupyter, MLflow, Agile/Scrum, MVC/MVT Architecture

EDUCATION

University Of Central Missouri — Warrensburg, Missouri
Master of Science in Computer Science — August 2022 – May 2024

Vel Tech University, Chennai — Tamil Nadu, India
Bachelor of Engineering in Electronics & Communication Engineering — July 2016 – May 2020

PROJECTS

Forecasting Agricultural Yield Trends using Time Series Analysis
Objective: Analyzing historical agricultural yield data using time series analysis to develop accurate forecasting models that inform decision-making and enhance agricultural planning and resource management.

- Utilized Python for time series analysis and regression to accurately forecast agricultural yields, presenting actionable insights through Tableau dashboards to optimize farming practices and productivity.
- Conducted EDA to identify trends, seasonal patterns, and correlations in the agricultural data.
- Tuned hyperparameters and performed cross-validation to optimize model performance.

Real-Time Data Processing System
Objective: Developing a scalable real-time data processing system using Apache Kafka and Spark Streaming to efficiently handle large volumes of streaming data while minimizing latency and ensuring data reliability.

- Engineered and implemented a real-time data processing system using Apache Kafka for reliable message queuing and
- Apache Spark Streaming for efficient data processing, enabling the handling of large volumes of streaming data.
- Developed data ingestion pipelines with Kafka Connect to stream data from various sources, transforming it in real-time using Spark Streaming to ensure timely and relevant insights.
- Managed the deployment of the system on cloud platforms, ensuring scalability and reliability, and performed regular maintenance to keep the system up-to-date and optimized.

CERTIFICATIONS

AWS Certified Solutions Architect Associate
AWS Certified Cloud Practitioner
Azure Data Scientist associate
Azure AI Fundamentals