Sumit Singh Patel

+91 6283095791 | sumitsnghpatel@gmail.com | LinkedIn | GitHub | Bangalore, KA

Summary

Machine Learning Engineer with 2.5+ years of experience in cloud-native, event-driven microservices and CI/CD. Expertise spans LLM integration, GenAI applications, MLOps pipelines, and Data Engineering, with a proven ability to lead cross-functional collaboration between data science and product teams to architect and deploy scalable AI products.

Work Experience

Honeywell Technologies Pvt. Ltd. **August 2023 – Present**
Data Engineer 2 *Bangalore, India*

- Developed and deployed predictive machine learning algorithms for smart building automation, reducing site onboarding time by 60-70% through automated asset classification using supervised learning techniques
- Built end-to-end ML pipelines using Python, FastAPI, and Databricks for model training, versioning, deployment, and real-time inference serving with sub-second latency requirements
- Architected MLOps infrastructure on Azure Kubernetes Service (AKS) with automated CI/CD pipelines, model monitoring, and A/B testing frameworks, improving incident resolution by 20%
- Optimized PostgreSQL databases for ML feature stores through advanced indexing and partitioning strategies, reducing query latency by 95% (from 60+ seconds to 3 seconds) for 300,000+ data points
- Implemented real-time ETL pipelines for IoT sensor data processing, enabling predictive maintenance models with anomaly detection capabilities using time-series analysis

Honeywell Technologies Pvt. Ltd. **May 2022 – July 2022**
ML Intern (PPO offered) *Bangalore, India*

- Implemented and compared various object detection algorithms including YOLO and Faster R-CNN for real-time helipad detection in video feeds
- Optimized detection frameworks to assist autonomous take-off and landing scenarios, evaluating performance metrics such as accuracy, precision, and inference speed

ADOPT Labs, IIT Madras — Dr. Palaniappan Ramu **May 2021 – July 2021**
AI Research Intern *Chennai, India*

- Developed time-series forecasting models using LSTM networks and ARIMA for trend analysis and pattern recognition in marketing analytics data
- Implemented Self-Organizing Maps (SOMs) and clustering algorithms for feature extraction and motif analysis in high-dimensional sequential data

Education

Indian Institute of Technology Madras **August 2019 – May 2023**
Bachelor of Technology, CGPA: 7.1/10 *Chennai, India*

Technical Skills

ML & Deep Learning: Pytorch, Tensorflow, SkLearn, LangGraph, LangChain, Ollama
Programming Languages: Python, C++, SQL, Java, Javascript
Cloud & Infrastructure: Azure, Databricks, Kubernetes (AKS), Docker
MLOps & LLMOps: Model Experiment Pipeline, MLFlow, Kubeflow, K-Serve, vLLM
DevOps & Tools: CI/CD (Bamboo, Octopus, Github Actions), Git, Postman, Coverity, Black-Duck
Backend & Databases: FastAPI, Postgres, pgVector, ETL pipeline development

Projects

Agentic Research Tool | *LangGraph, LangChain, Python, Bright Data, OpenAI*

- Built a multi-step AI research agent using LangGraph for graph-based workflow orchestration, enabling autonomous web research beyond basic prompt-response interactions
- Integrated Bright Data's SERP API and Web Unlocker for real-time web searching across Google, Bing, and Reddit with automatic anti-bot bypass for reliable data extraction
- Implemented stateful agent architecture with conditional edges, enabling dynamic tool selection and structured report generation in Markdown and JSON formats

Semantic Git-Keeper | *LangChain, FastAPI, Ollama, Python, Tree-sitter*

- Built a RAG system indexing Git history using Abstract Syntax Tree based chunking via tree-sitter
- Implemented hybrid search (vector + metadata filtering by author/file/function)for queries against codebase evolution
- Developed a diff-processing pipeline using local LLMs (DeepSeek-Coder via Ollama) to summarize commit intent
- Engineered co-change relationship graphs using NetworkX, establishing GraphRAG foundations for code archaeology