Alipt Shrivastava
Gurugram, Haryana

AI/ML Developer

+91-8349197741 | aliptassault@gmail.com LinkedIn | GitHub

Machine Learning Engineer with expertise in Python, Terraform, classification models, Conversational AI, Retrieval-Augmented Generation (RAG), and NLP. Experienced in LLM fine-tuning(PEFT, SFT), LangChain orchestration, lang graph, vector search and prompt engineering. Strong Devops,MLOPS skills, including AWS (SageMaker, Bedrock), Docker, CI/CD, and cloud-based AI deployments. Passionate about designing intelligent AI systems for automation, security, and real-time decision-making.

Skills & Technologies

Programming Languages & Frameworks: Python, LangChain, TensorFlow, PyTorch, Keras, FastAPI, NumPy, Pandas, scikit-learn
Machine Learning & AI: Supervised & Unsupervised Learning, Deep Learning, NLP, Generative AI, Large Language Models (LLMs), Langchain, Retrieval-Augmented Generation (RAG), Computer Vision, EDA, Feature Engineering, vector search databases.
Cloud & MLOps: Azure, AWS (Lambda, ECS, API Gateway, S3, DynamoDB, RDS, Bedrock, SageMaker), Terraform, Docker, CI/CD,
Databases: MySQL, PostgreSQL, NoSQL (DynamoDB)
DevOps & Tools: Git, Jira, Concourse CI/CD, Jenkins

Machine Learning Engineer | Fidelity International (2022 - Present)

1. Firewall Automation System
Implemented a **Random Forest Classifier** model to classify network risk zones, improving security accuracy to 94% and created a **xgboost** model to recommend firewall deployment with a 96% accuracy rate.
Automated firewall rule generation, reducing manual effort by 80% and deployment time by 60%.
Integrated continuous monitoring & feedback loops, retraining models for adaptive security enhancements.

2. Conversational AI-Powered Search & RAG Chatbot
Developed an advanced multi-agent chatbot system leveraging LangChain and multiple state-of-the-art LLMs (GPT-4, Claude 3.5 Sonnet, Gemma2-9b-IT, and Mistral) for robust, general-purpose information retrieval.
Integrated **RAG-based Q&A** using vector databases including **FAISS, Pinecone, Chroma, and Astra DB** to support context-aware, document-driven conversations across enterprise knowledge bases.
Utilized **HuggingFace embeddings** and **RecursiveCharacterTextSplitter** for efficient chunking and semantic search optimization over a large corpus of enterprise documents.
Designed and deployed intelligent multi-agent workflows using **LangChain Agents** (Zero-Shot ReAct, Few-Shot CoT) and **agentic orchestration frameworks** like **LangGraph** and **CrewAI** to facilitate modular task delegation and dynamic tool usage.

3. CISCO-ACI Network Configuration
Built a software tool for automating CISCO-ACI switch configurations, enhancing network efficiency by 30%.

4. Automated External DNS Management
Created an automation solution for DNS record creation, reducing errors and improving accuracy by 95%.

Key Projects:

1. Stock prediction model based on time-series financial data
Developed a stock market prediction model using **LSTM-RNN**, achieving **MSE = 0.032**.
Utilized genetic algorithm for hyperparameter tuning and Pearson's Correlation Coefficient for feature selection
Lead data preprocessing & feature engineering, ensuring improved model efficiency.

2. Emotion Detection using Deep Learning
Developed a CNN-based model for human emotion recognition, enhancing classification accuracy.

Education

Madhav Institute of Technology & Science, Gwalior (2018 - 2022)
Degree: Bachelor of Technology (B.Tech.)   CGPA: 7.84
Relevant Coursework: Machine Learning, AI, Cloud Computing, OOPS, DBMS, Data Structures & Algorithms
Certifications & Achievements
AWS Certified Cloud Practitioner,
Qualified GATE 2022,
Machine Learning Bootcamp – iNeuron,
Generative AI with langchain and huggingface – krishai technologies private limited