

Alipt Shrivastava
Gurugram, Haryana

AI/ML Developer
+91-8349197741 | aliptassault@gmail.com [LinkedIn](#) | [GitHub](#)

Machine Learning Engineer with expertise in **Conversational AI, Retrieval-Augmented Generation (RAG), and NLP**. Experienced in **LLM fine-tuning (PEFT, SFT), LangChain orchestration, lang graph, vector search (FAISS, Pinecone, Chroma, Astra DB), and prompt engineering**. Strong MLOps skills, including **AWS (SageMaker, Bedrock), Terraform, 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, TensorFlow, PyTorch, Keras, FastAPI, NumPy, Pandas

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. GPU acceleration, OpenCV, scikit-learn

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 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. CISCO-ACI Network Configuration

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

3. Automated External DNS Management

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

4. Conversational AI-Powered Search & RAG Chatbot

Developed an advanced **Streamlit**-based chatbot combining **LangChain** and **Groq API**.

Integrated **Arxiv**, **Wikipedia**, and **DuckDuckGo** search tools for real-time information retrieval.

Enabled **RAG**-based Q&A with (**FAISS, Pinecone, Chroma, Astra DB**) **vector search** for document-based interactions.

Implemented **HuggingFace** embeddings and **RecursiveCharacterTextSplitter** for optimized retrieval.

Designed and deployed using **LangChain Agents** (Zero-Shot React, Few-shot CoT) for intelligent query resolution using agentic workflows with tools like Langgraph, CrewAI.

Utilized LLMs (**Llama3-8b-8192, Gemma2-9b-It, Mistral, GPT**) for natural language processing and conversational AI.

Implemented session-aware chat memory with **RunnableWithMessageHistory** for persistent interactions.

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**.

Lead data preprocessing & feature engineering, ensuring improved model efficiency.

2. Forest Fire Prediction Model

Built a Random Forest Classifier to predict forest fires with 94.79% accuracy.

3. 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, GATE 2022 Qualified,

Machine Learning Bootcamp – iNeuron, Generative AI with langchain and huggingface