**Name:** Arjun Verma  
**Email:** arjun.verma@gmail.com  
**Phone:** +91-9876543210  
**LinkedIn:** linkedin.com/in/arjun-verma-ai  
**GitHub:** github.com/arjun-verma-ai

**Objective**

AI Engineer with 1-3 years of experience specializing in Generative AI, NLP, and Deep Learning. Skilled in developing and deploying AI solutions using LLMs, RAG, and fine-tuned models. Passionate about building scalable AI-driven applications and enhancing intelligent automation.

**Education**

**Bachelor of Technology (B.Tech) in Computer Science & Engineering**  
XYZ University, India | 2021

**Skills**

* **Programming:** Python, TensorFlow, PyTorch, Hugging Face, LangChain
* **AI/ML:** Generative AI, LLMs (GPT, BERT, T5, BART), RAG, Fine-Tuning, Transfer Learning
* **NLP:** Tokenization, Named Entity Recognition (NER), Text Summarization, Sentiment Analysis
* **Data Engineering:** Apache Spark, Hadoop, MongoDB, PostgreSQL, Kafka, Apache Flink
* **Cloud & Deployment:** AWS, GCP, Docker, Kubernetes, FastAPI, Streamlit
* **MLOps & Vector Search:** MLflow, Pinecone, FAISS, Weaviate, Elasticsearch

**Professional Experience**

**AI Engineer – ABC AI Solutions, Bangalore *(July 2022 – Present)***

* Developed **fine-tuned LLMs** for document summarization and chatbots using **T5/BART**.
* Implemented **Retrieval-Augmented Generation (RAG) with FAISS** for AI-powered search solutions.
* Designed a **real-time fraud detection pipeline** using **Kafka, Spark Streaming, and MongoDB**.
* Deployed AI models on **AWS Lambda & GCP Cloud Run**, optimizing cost and scalability.
* Integrated **LLMs with Function Calling APIs** to automate business operations.

**Data Scientist – XYZ Analytics, Hyderabad *(Jan 2021 – June 2022)***

* Built **sentiment analysis and NLP pipelines** for customer behavior prediction.
* Developed **multi-modal search engines** using **Elasticsearch and OpenAI embeddings**.
* Created **real-time dashboards with Tableau/Power BI** for AI-driven insights.
* Optimized **ML pipelines using MLflow and Docker for efficient model deployment**.

**Real-Time Projects**

**1. AI-Powered Resume Screening System**

* Developed a **resume screening AI** using **GPT-4 & BERT embeddings**.
* Implemented **vector search (Pinecone)** for efficient retrieval of resumes.
* Deployed using **FastAPI & Streamlit** on **AWS Lambda** for seamless automation.

**2. Multi-Modal AI Search Engine**

* Built a **search engine** integrating **text and image embeddings** using **CLIP & OpenAI embeddings**.
* Used **FAISS & Elasticsearch** for high-speed retrieval across structured/unstructured data.
* Designed a **Kafka-based real-time data ingestion pipeline** for dynamic indexing.

**3. AI-Powered Personal Assistant**

* Designed an **AI Assistant** that integrates **voice, document retrieval, and task automation**.
* Used **Whisper for Speech-to-Text** and **TTS models for voice responses**.
* Deployed on **GCP Cloud Run** with vector search for personalized recommendations.

**4. Financial Fraud Detection System**

* Built an **AI-driven fraud detection model** using **LSTMs and Isolation Forest**.
* Implemented **Kafka-based event streaming** for real-time fraud alerts.
* Deployed using **Databricks, AWS Lambda, and SNS for instant notifications**.

**Certifications**

* **Deep Learning Specialization – Coursera (Andrew Ng)**
* **Generative AI with Large Language Models – DeepLearning.AI**
* **AWS Machine Learning Specialty Certification (AWS-MLS)**

**Academic Projects**

* **Text Summarization using Fine-tuned BART/T5** – Optimized for document summarization.
* **Conversational AI Chatbot using OpenAI GPT-4 API** – Built an LLM chatbot with memory and context awareness.
* **Stock Market Prediction using Transformers** – Developed AI models for time-series forecasting.

**Achievements**

* **Winner of Hackathon on AI-powered Resume Parsing (IIT Bombay, 2023)**
* Published research on **RAG-based AI Search** in **IEEE Conference (2023)**
* **Top 5%** in **Kaggle NLP Competitions** for sentiment analysis and text classification.