

PROFESSIONAL SUMMARY

Generative AI Engineer skilled in developing and fine-tuning LLMs, Retrieval-Augmented Generation (RAG), and optimizing Transformer-based architectures. Experienced in deploying AI solutions using LangChain, Hugging Face, and TensorFlow. Passionate about leveraging AI for real-world applications with scalable and efficient models.

WORK EXPERIENCE

Data Quality Analyst — Arihant Technologies Ajmer, Rajasthan
Jan. '24 – Present

- Developed end-to-end Generative AI applications, integrating large language models with vector databases using Retrieval-Augmented Generation (RAG) pipelines for context-aware, domain-specific chatbots.
- Designed scalable backend workflows for dynamic data ingestion, embedding generation, and semantic search, ensuring real-time, accurate responses in production-grade AI assistants.
- Led a team of 6 annotators, improving dataset quality by 32%, leading to a 14% increase in model accuracy.
- Optimized annotation workflows, improving data integrity and consistency.
- Collaborated with the ML team to refine data requirements, enhancing model performance.

Technical Support Engineer — Talview Bengaluru, Karnataka
Aug. '21 – Dec. '23

- Provided technical support for the Talview Platform, ensuring smooth client operations.
- Resolved 500+ technical issues monthly using Insomnia, GitHub and Jira, reducing downtime by 25%.
- Collaborated with engineering teams to improve platform performance and user experience.

EDUCATION

Bachelor of Technology — Government Engineering College Ajmer Ajmer, Rajasthan
Jul. '12 – Jan. '18

- Studied core subjects including Data Structures, Algorithms, Operating Systems, Computer Networks, Database Management Systems, and Web Technologies.
- Completed academic projects focused on software development, system design, and IT infrastructure.
- Gained practical experience through labs, coding assignments, and mini-projects using languages like C++, Java, and Python.
- Participated in technical seminars, coding competitions, and departmental events to enhance technical and teamwork skills.

SKILLS

- **Machine Learning & AI:** Deep Learning, Transformers, Fine-Tuning, Recommendation Systems, Evaluation Techniques, Agentic AI.
- **NLP (Natural Language Processing) & Generative AI:** RAG, Prompt Engineering, LLMs (Large Language Models), Embeddings, GANs, Hugging Face, Ollama.
- **Computer Vision & Time Series:** OpenCV, Stable Diffusion, Time Series Analysis.
- **Data Processing & Analytics:** Scikit-Learn, SciPy, Data Visualization (Tableau, Power BI).
- **Frameworks & Libraries:** Pandas, NumPy, Matplotlib, Seaborn, Plotly, TensorFlow, Keras, Nltk, Spacy, LangChain, Gensim.
- **Vector Databases & Retrieval:** ChromaDB, FAISS, Pinecone, Astra DB.
- **Deployment & MLOps:** MLflow, AWS Ec2, AWS Bedrock, Streamlit, Docker, Kubernetes, Git, GitHub, Google Colab.
- **Programming & Databases:** Python, SQL, MongoDB.

PROJECTS

Financial AI Assistant (Python, LangChain, YFinance, DuckDuckGo API)

- Developed an AI-powered financial recommendation system using multi-agent models.
- Integrated real-time financial news and stock data, improving recommendation accuracy.

Calorie Counter – Large Image Model (LIM) (Gemini, OpenCV)

- Built an AI-based food image recognition system to estimate calorie values.
- Deployed using Gemini API, achieving 90% accuracy in calorie estimation.

Conversational AI Bots (Google Gemini, Streamlit)

- Developed a Data Science Instructor bot with memory retention for enhanced user interaction.
- Implemented custom AI pipelines, leading to 20% better response accuracy.

RAG-Based Q&A Bots (ChromaDB, LangChain)

- Designed and deployed two knowledge-based AI assistants querying Bhagavad Gita & Ramayana.
- Improved retrieval efficiency by 40% using optimized vector search techniques.

Medical Test Results Classification (Scikit-Learn, Pandas, Machine Learning)

- Built an ML classification model to categorize medical test results into Normal, Abnormal, and Inconclusive.
- Trained using structured patient data, achieving high classification accuracy in clinical test datasets.
- Implemented data preprocessing techniques to handle missing values and outliers, improving model reliability.