Vishal Shivnani

Email : vishal1594@outlook.com
Mobile : +91-8619748600

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.

EXPERIENCE

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

- Led a team of annotators to ensure high-quality labeled datasets for ML model development.
- Optimized annotation workflows, improving data integrity and consistency.
- Collaborated with the ML team to refine data requirements, enhancing model performance.

Technical Support Engineer — Deep Learning Engineer Bengaluru, Karnataka
Aug. '18 – Dec. '23

- Provided technical support for the Talview Platform, ensuring smooth client operations.
- Diagnosed and resolved issues using Insomnia, Jira, and advanced MS Excel.
- Collaborated with engineering teams to improve platform performance and user experience.

SKILLS

- **Machine Learning & AI**: Deep Learning, Transformers, Fine-Tuning, Recommendation Systems, Evaluation Techniques, Agentic AI.
- **NLP & Generative AI**: RAG, Prompt Engineering, 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:** AWS Bedrock, Streamlit, Docker, 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.

YouTube Transcript Summarizer (Python, YouTube Transcript API)

- Developed an AI-powered summarization model that extracts and condenses transcripts from YouTube videos.
- Utilized YouTube Transcript API and NLP techniques to generate concise, context-aware summaries, improving readability and efficiency.
- Deployed as a web-based tool, enhancing accessibility for content consumers.

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.