Deepak Pawar

📞 +91-8770406446   👤 deepakpwiitk@gmail.com   in - LinkedIn   ⭘ - GitHub

Education
Bachelor of Technology: Civil Engineering, 2013-2017
Indian Institute of Technology, Kanpur

Professional Summary
Results-oriented **Data Scientist** with over **2.5 years** of experience in leveraging **GenAI**, **Deep Learning**, and **Machine Learning** techniques to solve complex **AI Agent** and **NLP** challenges, enhance generative analytics, and deliver actionable insights that fuel business growth. Skilled in **MLOps tools**, developing, and deploying end-to-end NLP solutions within **CI/CD pipelines** for seamless integration and continuous delivery of models and applications.

Skills
- Languages – **Python**
- AI Agents – **SQL AI Agent, Conversational AI Agent and Chatbots, Agentic RAG, Agentic AI, MCP Integration**
- GenAI – **LangChain, LangGraph, LLM, LLAMA, RAG, BERT, OpenAI, Ollama, GROQ, Transformers, Hugging face**
- DL & ML – **ANN, RNN, LSTM, GRU, Regression, Ensemble Methods, Boosting Methods, Hypothesis Testing**
- Data Engineering - **ETL Pipeline, Data Cleaning, Feature Engineering, EDA, Data Pipeline**
- Clouds – **Amazon Web Service [Lambda, Bedrock, Sagemaker, EC2, S3], Microsoft Azure**
- Databases – **MySQL, MongoDB, PostgreSQL**
- Libraries – **TensorFlow, Keras, PyTorch, NLTK, Spacy, Scikit-Learn, Pandas, NumPy, Matplotlib, Seaborn**
- Web Development - **Flask, Streamlit, HTML, CSS**
- MLOps Tools – **Docker, Git, GitHub Action, DVC, mlflow, Apache Airflow, Dagshub, Grafana, Notebook, Colab**

Work Experiences
Data Scientist, at ABM   [*Nov'22-Present*]
- Developed and deployed an automated **SQL AI Agent** system leveraging **LangChain**, **Agent Class**, and **Streamlit**, utilizing **LangChain Chaining** techniques to generate, retrieve, and execute accurate **SQL** queries from **natural language** and **audio inputs**. The system allows users to view results in a simple text format.
- Designed a robust Document **Q&A chatbot** using **LangChain, ChromaDB,** and **Streamlit** utilizing **Retrieval Augmented Generation (RAG)** techniques to extract relevant information from **PDFs** by leveraging their text, summaries, metadata. Utilized **HuggingFace** embedding and **OpenAI LLM models** for efficient processing.
- Developed a **Machine learning** model to convert captcha images into text, achieving an accuracy of **99.95%**. The implementation of this project reduced the time required for the voter roll scraping from **5 days** to **5 hours**.
- Developed and fine-tuned a **BERT** based model for **text classification** and **named entity recognition (NER)** on a custom dataset, achieving high accuracy and domain-specific performance. Optimized with advanced techniques and deployed the model for real-time NLP applications.

Junior Commissioned Officer, at Border Road Organization   [*2020-2022*]
- Served as Junior Commissioned Officer in **BRO** for improving the road connectivity of inaccessible and remote locations of Sikkim while withstanding the adverse weather and terrain challenge.
- Designed road pavement structure, culvert and bridges by using software such as **AutoCAD**, **IITPAVE** and **QGIS.**

Prepared for Civil Services Examination   [*2017-2020*]

Projects
- **SQL AI Agent**, code - ⭘ , video - ▶
  - **Aim :-** Developed an AI-powered **SQL Agent** leveraging **OpenAI GPT-4.1** and **Whisper** to convert text and voice inputs into precise SQL queries, enabling non-technical users to interact seamlessly with **relational databases**.
  - Integrated **OpenAI Whisper** for real-time speech-to-text transcription, enabling voice-driven query generation, which is further processed by **OpenAI GPT-4.1** to generate precise **SQL commands**.
  - Engineered a modular, multi-agent architecture using **LangChain's agent** and chaining framework to handle query parsing, validation, execution, and result summarization in a structured pipeline.
  - Developed an intuitive **Streamlit UI** to facilitate seamless user interaction with the SQL Agent, supporting both voice and text inputs, and dynamic result visualization.
- **Conversational Chatbot (RAG)**, code - ⭘
  - **Aim :-** Developed a **Conversational RAG Chatbot** leveraging **OpenAI GPT-4.0** model to enable intelligent question answering over user-uploaded PDF documents using semantic search and vector similarity.
  - Implemented document ingestion pipeline that parses, chunks, and indexes PDF files into a **Chroma** vector database using embedding models for efficient retrieval. Also, Integrated **LangChain's conversational memory modules** to maintain and utilize chat history, enabling context-aware multi-turn conversations.
  - Built a **Streamlit**-based user interface supporting PDF uploads, real-time chat interaction, and visual feedback.

Interests - GenAI, NLP, LLM, AI Agent, Chatbot, RAG, Deep Learning, Data Science, Machine Learning etc.