Hricha Jaiswal

Machine Learning Engineer
Think Dignity International, India.
Portfolio: https://hricha-jaiswal-portfolio.onrender.com/

+91-9140308424
hrichajaiswal@gmail.com
https://github.com/Hricha-1511
https://www.linkedin.com/in/hricha-jaiswal/

Summary

A highly skilled **Machine Learning Engineer** with a Master's degree in Mathematics and Computing from **IIT ISM Dhanbad**, bringing robust Python programming expertise and a deep understanding of machine learning algorithms. Experienced in developing innovative AI solutions, I have a strong background in text-to-SQL models, RAG, and automation workflows using FastAPI, Flask, PostgreSQL, make and n8n. My proactive approach in staying updated with AI advancements enables me to tackle complex problems, deliver tailored solutions, and contribute effectively to research and development efforts. I am passionate about leveraging my technical skills to enhance AI capabilities.

Education

•Master of Science in Mathematics and Computing *2021-23*
Indian Institute of Technology, ISM, Dhanbad CGPA: 7.49

•Bachelor of Science in Mathematics (Hons.) *2017-20*
Banaras Hindu University, Varanasi CGPA: 7.77

Technical Skills and Interests

Languages and Softwares: Python, Power BI, SQL, PostgreSQL, Jupyter Notebook, DAX, C++, Botpress, Vanna
Libraries: Autogen, pyautogen, streamlit, PyPDF2, gtts, speech recognition, Numpy, Pandas, Matplotlib, Seaborn, sklearn, nltk, openai, fastapi, flask
Areas of Interest: Machine Learning, Data Science, Data Analytics, Research
Soft Skills: Problem Solving, Self-learning, Presentation, Adaptability, Verbal and Written Communication

Experience

•Machine Learning Engineer (FT) *May 2024 - Oct 2024*
TDI-India Remote

- Spearheaded a project focused on improving the precision and reliability of a model that translates natural language queries into SQL **(text2SQL)**.
- **Implemented FastAPI Backend:** Designed and deployed a **FastAPI** backend to process and deliver query outputs efficiently, ensuring seamless integration with the frontend.
- Utilized **PostgreSQL** to manage and retrieve data for a specialized US market database, ensuring high accuracy.
- Regularly engaged with **clients** to understand their needs, gather feedback, and provide tailored solutions, enhancing the overall product experience.
- Contributed to the **Research and Development** team by conducting research on Prompting Techniques, Retrieval-Augmented Generation (RAG), cosine vectors, and ambiguity resolution.
- Developed a product using **Autogen** for conversations with LLMs, integrating **SpeechRecognition** for voice-to-text conversion and **gTTS** for audio response generation.
- Created automation workflows using **n8n** with **Pinecone** vector database for chatbot responses to user queries.
- Automated property data processing using **n8n** and **Cloudinary** to extract images and text from files and apply watermarks for specific use cases.
- Developed an **n8n** chatbot for real-time **Google Scholar** searches, extracting summaries of the top 5 research articles and providing concise outputs.
- Built a multimodal advanced RAG-based chatbot using **llama-index** and **llama-parse** for PDF data extraction, capable of providing answers from images within PDFs, and deployed it on a server using flask.
- Worked on **Odoo** server automations for CRM, sales, email workflows, and other business processes.
- Designed multi-step **email automation** in Odoo, triggering actions based on interactions like email or link opens.
- Implemented automation in **make.com** to create LinkedIn posts with images based on new RSS feeds.
- Automated invoice processing using **n8n**, capturing email triggers and storing summarized data in databases.

•Data Science Intern *Feb 2024 - Apr 2024*
TDI-India Remote

- Utilized **Autogen, LLM, and RAG** technologies to enhance data analysis, driving more efficient and accurate insights.
- Applied **advanced mathematical concepts** to improve the interpretative capabilities of LLM and RAG models, ensuring more precise predictions and analyses.
- Synthesized mathematical techniques with AI methodologies to solve complex challenges, refining algorithms and spearheading AI innovation.

•Data Science Intern *Nov 2023 - January 2024*
Success Ladder Technology Pvt. Ltd. Remote

- Developed a specialized **NER** system tailored for the vehicle domain using OpenAI's GPT model.
- Implemented **few-shot learning techniques** to fine-tune the OpenAI GPT model in the vehicle domain.
- I conducted comprehensive testing and evaluation of the model's performance with real-world user queries, and integrated the model into **Elasticsearch**.