Debajyoti majee

Adm. No. 23JE0291 8926232518
majeedebajyoti2004@gmail.com linkedin/Debajyoti Majee github/Debajyoti2004

Education

IIT (ISM) Dhanbad Expected May 2028
Integrated M.Tech in Mathematics and Computing (GPA: 8.99 / 10.00) *Dhanbad, Jharkhand*

- **Relevant Coursework:** Data Structures and Algorithms (C++) , Prob & Stat, Linear Algebra , Real Analysis, Statistical Inference , Discrete Mathematics

Projects

Resume Builder Project | *RAG, LangChain, Gemini 2.0 Flash LLM, Streamlit, FAISS, Web Search* GitHub Link

- Developed an AI-powered resume builder leveraging RAG to retrieve and generate personalized content by combining real-time web search results with stored resume data.
- Integrated LangChain to orchestrate the RAG pipeline — fetching relevant information from web search, processing through Gemini 2.0 Flash LLM, and dynamically crafting tailored resume sections.
- Implemented a transfer query mechanism, automatically refining queries if the retrieved data is irrelevant, ensuring high-accuracy content generation.
- Utilized FAISS for vector-based text retrieval, optimizing response time by 40%, ensuring efficient access to user data and job-specific keywords.
- Built an interactive interface with Streamlit, allowing seamless input of career details and real-time resume preview.

Document Text Extractor | *Flask, PyTesseract OCR, spaCy NER, BIO Tagging, Pandas* GitHub Link

- Developed a document text extraction tool using PyTesseract OCR to accurately retrieve text from scanned documents and images.
- Implemented spaCy NER with BIO tagging to identify and classify key entities (e.g., names, dates, and addresses) and present them in a structured table format.
- Utilized Pandas to process and organize extracted data, enabling seamless manipulation and conversion into CSV or JSON formats.
- Built an interactive user interface with Flask, allowing users to upload documents and view extracted text and entity tags in real-time.

Dialogue Summarizer | *LoRA PEFT, FLAN-T5, PPO Optimization, Hugging Face, Numpy, RL* GitHub Link

- Built a dialogue summarization model using LoRA PEFT fine-tuning on FLAN-T5 LLM to generate concise and accurate summaries.
- Applied PPO optimization to mitigate harmful generation, ensuring outputs remain relevant and safe.
- Integrated a toxicity prediction model during text generation, feeding toxicity scores as rewards to the PPO optimizer to guide safer and more responsible outputs.
- Leveraged Hugging Face libraries for seamless training and inference, enhancing model deployment efficiency.
- Processed and manipulated training data using Numpy, ensuring efficient handling of tokenized inputs and model outputs.

Technical Skills

Languages: Python, JavaScript, C++, SQL
Web Technologies: HTML5, CSS3, React.js
Technologies: PyTorch, TensorFlow, LangChain, RAG, FAISS, Flask, Streamlit, Numpy, Pandas, Sklearn, Hugging Face, Gemini 2.0 Flash LLM
Concepts: Reinforcement Learning, Fine-tuning (LoRA PEFT), RAG Pipelines, Neural Networks, Machine Learning, Deep Learning/AI, Natural Language Processing, Operating Systems

Achievements

- Secured AIR 3019 in JEE Mains 2023 with a 99.74 percentile.
- Achieved AIR 5502 in JEE Advanced 2023.