Intel Unnati

Team sparks

Problem statement -[PS16]

Running GenAI on Intel AI Laptops and Simple LLM Inference on CPU and fine-tuning of LLM Models using Intel® OpenVINO™

Unique Solution -Optimized Medical Chat bot

Performance Optimization with OpenVINO:

- **Enhanced Efficiency:** Utilizes OpenVINO to optimize the Llama 3B model, ensuring it runs efficiently on both CPUs and GPUs.
- **Faster Response Times:** Achieves low latency and quick responses, crucial for real-time interaction.
- **Precision Control:** Allows fine-tuning of performance to balance speed and accuracy according to specific needs.

Advanced Medical Knowledge:

- **Extensive Training:** Trained on a comprehensive corpus of medical literature to ensure accuracy and reliability.
- **Contextual Understanding:** Capable of understanding and responding to complex medical queries with detailed information about symptoms, treatments, and procedures.

Features Offered

Multilingual Support:

- **Inclusive Language Options:** Supports multiple languages, including English and Hindi, making it accessible to a diverse user base.
- **Language-Specific Examples:** Provides examples and responses tailored to different languages, enhancing user experience for non-English speakers.

User-Centric Customization:

- **Adjustable Parameters:** Users can customize interaction settings such as temperature, top-p, top-k, and repetition penalty, allowing control over the chatbot's response diversity and creativity.
- **Intuitive Interface:** Built with Gradio, providing an easy-to-use and interactive interface for users to engage with the chatbot.

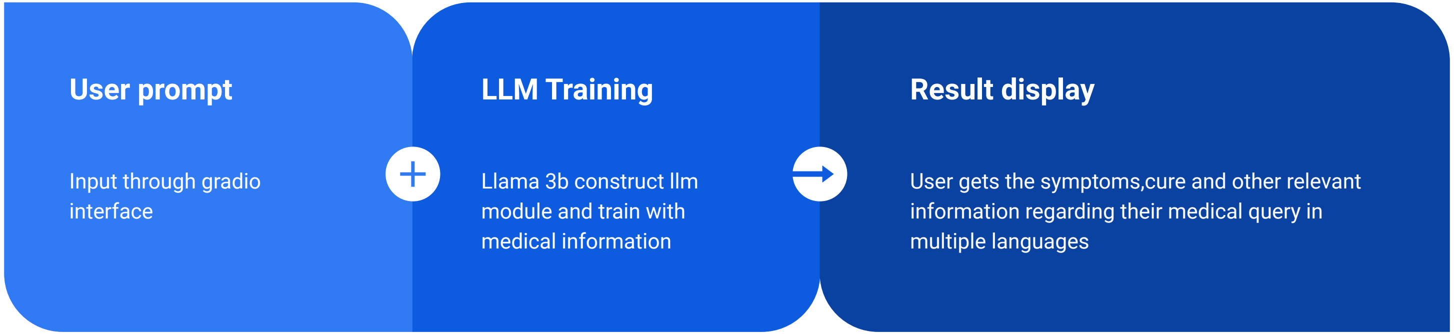
Real-Time Interaction:

- **Seamless Conversations:** Utilizes TextIteratorStreamer and threading to handle real-time text generation, ensuring smooth and interactive conversation flows.
- **Scalable Architecture:** Designed to be scalable, suitable for various deployment environments from personal use to integration within healthcare systems.

Educational and Supportive Roles:

- **Healthcare Assistant:** Acts as a supplementary tool for healthcare professionals, aiding in patient education and providing preliminary advice.
- **Educational Resource:** Serves as a learning tool for medical students and professionals, offering detailed explanations of medical conditions and procedures.

Process Flow



User prompt

Input through gradio interface

LLM Training

Llama 3b construct llm module and train with medical information

Result display

User gets the symptoms,cure and other relevant information regarding their medical query in multiple languages

Technologies used

- **Intel OpenVino**
- **Hugging Face**
- **Gradio**
- **Llama 3b**

Team members and Contribution

Allen Bijo T

- **Model Integration:** Implemented the Llama 3B model with OpenVINO for optimized performance.
- **Medical Data Training:** Ensured the model was trained on extensive medical literature for accuracy.

Shreya Roshan

- **Multilingual Support:** Developed and tested the chatbot's multilingual capabilities.
- **User Interface Design:** Created an intuitive and interactive interface using Gradio.

Vrushika SunilKumar Modi

- **Customization Features:** Implemented adjustable parameters for user customization.
- **Real-Time Interaction:** Managed real-time text generation and threading for seamless conversations.

Conclusion

MeduChat is a groundbreaking AI-driven medical chatbot that leverages the Llama 3B model optimized with OpenVINO for enhanced performance and efficiency. It offers:

- **Optimized Performance:** Fast and efficient real-time responses.
- **Accurate Medical Information:** Reliable answers from extensive medical training.
- **Multilingual Support:** Accessible in multiple languages, including English and Hindi.
- **User Customization:** Adjustable parameters for tailored responses.
- **Real-Time Interaction:** Smooth, interactive conversations.
- **Educational Resource:** Useful for both healthcare professionals and students.

MeduChat is a comprehensive solution that improves access to reliable medical information, making it a valuable tool in the healthcare sector.