

AN INTELLIGENT MULTI-AGENT SYSTEM FOR PAEDIATRIC ONCOLOGY PATIENT SUPPORT USING LARGE LANGUAGE MODELS

Project Reference No.: 49S_BE_2281

College : New Horizon College of Engineering, Bengaluru
Branch : Department of Artificial Intelligence and Machine Learning
Guide(s) : Prof. Ramyashree P M
Student(s) : Ms. Rajalakshmi R
 Ms. Sonal Shirke
 Mr. Kiruthik J

Keywords:

Mistral 7B, Agentic AI, LLM, Pediatric Oncology, Machine Learning

Introduction:

1. Pediatric oncology is a specialized area of medicine that focuses on the study, diagnosis, and treatment in children. Over 400,000 new patients are affected by childhood cancer each year, making it one of the major causes of mortality for children globally. Only a small portion of the almost 50,000 children who are diagnosed in India alone each year receive prompt treatment due to socioeconomic constraints, lack of awareness, delayed diagnosis, and poor infrastructure. These challenges sometimes result in poorer survival rates than in wealthy countries, highlighting the urgent need for innovative alternatives beyond pharmacological therapies.
2. AI solutions help in this domain by automating the data flow during various treatment procedures, learning from large amounts of pediatric cancer-related data and for personalized patient care. AI-powered conversational agents and chatbots are one of these technologies that can be used for tasks such as patient education, mental health assistance, symptom screening, and post-treatment follow-up. This advancement has been further expedited by the use of LLMs. Large volumes of accurate, pre-processed, clinical data are used to train LLMs. When refined, they can develop into extremely specialized instruments that can forecast possible adverse effects, provide tailored advice, and convert technical medical terms into understandable terms for young patients. Although LLMs for adult cancers exist, it is necessary to fine-tune LLMs specifically for child cancer, to ensure the models do not confuse adult and child cancers and perform tasks for pediatric cancer domain more effectively.
3. We are focused on addressing the fears and questions of child cancer patients and their parents/guardians. Our solution aims to make parents/guardians medically aware, well-informed, eliminate taboos and help them reduce their fear by providing them and child patients with patient care through the application, powered by ethical use of AI.

Objectives:

1. A chatbot (OncoBot) using GPT-4 for empathetic, child-friendly conversations regarding any questions a user might have.
2. Implementing a predictive model using Mistral 7B (CancerLLM) to anticipate symptoms and side effects specific to the user's cancer.
3. A multi-agent framework using OpenAI Swarm to seamlessly connect user data and all backend models.
4. Risk stratification model designed to analyze patient's demographic data and treatment plans to obtain success rate of said treatment.

Methodology:

The system is fundamentally built through the multi-agent framework. The features are as follows:

- 1) Collecting users' medical data from patient reports
Preprocess the sample reports and obtain important, relevant details and store them in a text file.
- 2) Handle risk stratification tasks using the Machine Learning model - DecisionTreeRegressor which is trained on synthetically developed pediatric cancer data that contains various features like demographics, subsequent malignancies, recurrences, other causes like cardiac and pulmonary.
The primary purpose is to produce a probabilistic value of success/failure rate of the patient's treatment. Hence, the user data is fed to the model which is saved using Pickle module and preprocessed similar to the training data.
- 3) Connect with Mistral7B, the CancerLLM, for symptoms and side effects of cancer.
The CancerLLM is trained on a custom dataset, and thus delivers strong performance while remaining lightweight and fast.
The model is loaded from HuggingFace in a quantised, memory-efficient format for the fine-tuning. A custom dataset was prepared by combining textual data from pediatric oncology case studies, patient symptom logs, caregiver reports, and publicly available medical literature. It includes the possible common symptoms of child cancer and data outlining the treatment stages, drug effects, and the side-effects.
The evaluation metrics used for the fine-tuned Mistral 7B model were ROUGE and BLEU.

Result and Conclusion:

The developed intelligent multi-agent system successfully integrates multiple AI components to provide comprehensive support for pediatric oncology patients. The chatbot (OncoBot) demonstrated the ability to engage in empathetic, child-friendly conversations while delivering medically relevant information. The fine-tuned Mistral 7B model (CancerLLM) effectively generated accurate predictions of symptoms and treatment side effects, achieving satisfactory performance based on BLEU and ROUGE evaluation

metrics. Additionally, the risk stratification module using the Random Forest Regressor produced reliable probabilistic estimates of treatment success based on patient-specific inputs. The seamless coordination of these modules through the multi-agent framework ensured efficient data flow and contextual response generation.

References:

- [1] M. Moor et al., “Foundation models for generalist medical artificial intelligence,” *Nature*, vol. 616, pp. 259–265, 2023. doi:10.1038/s41586-023-05881-4.
- [2] M. Li et al., “CancerLLM: A large language model in the cancer domain,” *arXiv preprint*, arXiv:2406.10459v1, 2024.
- [3] F. Jia et al., “OncoGPT: A medical conversational model tailored with oncology domain expertise on a large language model,” *arXiv preprint*, arXiv:2402.16810, 2024.
- [4] N. Mehandru et al., “Large language models as agents in the clinic,” *arXiv preprint*, arXiv:2309.10895, 2023.
- [5] S. Schmidgall et al., “AgentClinic: A multimodal agent benchmark to evaluate AI in simulated clinical environments,” *arXiv preprint*, arXiv:2405.07960, 2024. <https://doi.org/10.48550/arXiv.2405.07960>

Future Scope:

The future scope of this project includes:

1. To store user information in a secure database.
2. To process both images and textual patient reports.
3. To utilize the CancerLLM with larger dataset that consists of in-depth medical data and more cancer types.
4. To build a mobile application for better user experience and accessibility.