Developing LLM Agents for Resilient, Efficient, and Ethical Capacity Modelling in Health Care Provision

Abstract

Capacity modelling in health care provision is a systematic problem to assessing, planning, and optimising the allocation of resources, such as staff, equipment, and facilities, to meet patient demand and improve the resilience, efficiency, and ethics of health care services. It has been proved crucial, especially during pandemics.

Many objectives need to be considered in capacity modelling, such as service delivery efficiency, patient outcomes, and resilience, alongside many constraints, including limited resources, fluctuating patient demand, regulatory requirements, and ethics. It is an open problem to achieve the Pareto optimal solution of this multi-objective, constrained problem.

To address this issue, this project proposes to design an AI agent based on large language models (LLMs) for the capacity modelling problem. Our objectives include (1) designing model architectures and optimisation algorithms of LLM agents for capacity modelling in health care provision, (2) establishing mathematical guarantees of this LLM agent's generalisability, stability, algorithmic fairness, etc., and (3) developing a prototype system employing the developed algorithms for real-world applications.

1. Introduction

Capacity modelling in health care provision is a quantitative and analytical problem of assessing, planning, and optimising the resources, such as staff, equipment, and facilities, required to meet the health care needs of a patient cohort or the wider population. Capacity modelling helps decision-makers evaluate the impact of changes in resource allocation, operational processes, or patient flow, ultimately supporting strategic planning for multiple objectives bounded by constraints: (1) **resilience**: ensuring resilience in the face of public health crises, such as unexpected surges in demand, or unforeseen reduction in supplies, (2) **efficiency**: enhancing patient outcomes, minimising delays in care, reducing waiting times, and ensuring timely access to medical services, and (3) **ethics**: ensuring equitable resource allocation, transparency in decision-making, patient autonomy, etc. Capacity modelling is thus a multi-objective, constrained problem in health care provision. It remains an open problem to achieve the Pareto optimal solution of this.

This project proposes to design an AI agent based on large language models (LLMs)[1] to tackle the capacity modelling challenges in health care provision. The LLMs are pre-trained on massive language data, aiming to learn 'general' knowledge, and then fine-tuned for decision-making in capacity modelling, leveraging the knowledge learned in the pre-training stage. This

[1] Carta, T., Romac, C., Wolf, T., Lamprier, S., Sigaud, O. and Oudeyer, P.Y. Grounding large language models in interactive environments with online reinforcement learning. International Conference on Machine Learning 2023.

paradigm has seen preliminary successes in the route planning of autonomous vehicles. In summary, the objectives of the project are three-fold:

a. designing models and optimisation algorithms for LLM agents capable of solving the multi-objective, constrained optimisation problems induced by capacity modelling,
b. establishing the mathematical foundations underlying this LLM agent, from critical aspects of generalisability (ensuring that the model performs well in unseen data), regret cumulated alongside the time, convergence speed in training the LLM agent, stability of deployment in fluctuating environments, and algorithmic fairness ensuring equitable solutions across diverse patient groups. The theory will provide foundational guarantees for the agent's reliability and ethical implementation, and
c. developing a prototype system that integrates the designed algorithms, pursuing their applications in real-world health care provision.

This PhD project is in collaboration with Brigham and Women's Hospital, a top hospital in the United States, and a premier teaching hospital of Harvard Medical School. The Brigham side is led by Dr Lisa Lehmann, as a co-supervisor. She is the Director of the Center for Bioethics at the Brigham, an Associate Physician in the Jen Center for Primary Care of the Brigham, and an Associate Professor at Harvard. She was the Director of Bioethics and Trust at Google, and the Chief Medical Office of the U.S. Department of Veterans Affairs when she managed a multi-state integrated healthcare system comprised of 8 medical centres and 41 community-based outpatient clinics that serves over 260,000 Veterans.

2. Research Challenges

We identify four key challenges in realising the objectives:

a. how to design the LLM agents' model architectures and optimisation algorithms to be capable of capacity modelling in health care provision,
b. how to transfer the knowledge learned from pretraining LLMs to decision-making for capacity modelling in health care provision,
c. how to overcome the mathematical difficulties of establishing the theoretical foundations. It has been commonly accepted that deep learning lacks a solid theoretical foundation,
d. how to avoid issues of stability in fluctuating environments, algorithmic fairness for diverse patient groups, etc., and
e. how to integrate all algorithms into a capable prototype system that can be used in real-world scenarios.

3. Data & Methodology

Methodology: The research will develop a full suite of models and optimisation algorithms of LLM agents for capacity modelling in health care provision, with mathematical guarantees, which will be then integrated into a prototype system for real-world scenarios. It will be an interdisciplinary research project – related techniques including LLM, multi-objective constrained optimisation, reinforcement learning, algorithmic fairness, mechanism design, algorithmic game theory, learning theory, optimisation theory, etc., whilst domain knowledge in health care system, especially the capacity modelling of health care provision, is crucial.

Timetable:

a. literature review (months 1 to 6)
b. designing the model architectures and optimisation algorithms of an LLM agent for capacity modelling in health care provision (months 7 to 18)
c. mathematically analyse this LLM agent from aspects including generalisability, regret, convergence, algorithmic fairness, etc. (months 19 to 30)
d. developing a prototype system of employing developed algorithms (months 31 to 42)
e. preparing thesis (months 42 to 48)

Data: The experiments will be conducted in two stages: (1) verifying the developed algorithms in open-access data and simulation environments, and (2) deploying the developed algorithms to the external partner's data under ethical approval.

4. Responsible AI & Ethical Considerations

The team recognises that the algorithmic fairness of the LLM agents is critical in health care provision. Training LLM agents often relies on large datasets that may contain historical biases reflecting inequities in health care access, treatment, and outcomes. These biases can arise from a range of factors, including socioeconomic status, race, ethnicity, gender, geographic location, and disability. If not addressed, LLM agents may unintentionally perpetuate or even amplify these disparities by disproportionately allocating resources to populations or regions already favoured by existing health care systems, thereby marginalizing underserved groups.

Algorithmic fairness is in the central place of this project: (1) fairness is amongst the prioritised objectives and constraints in training LLM agents, aiming to control the risks from the first instance; (2) the project endeavours to establish mathematical guarantees of fairness, mitigating risks from theoretical stems; and (3) the external partner led by Dr Lehmann will make navigate the ethical considerations.

5. Expected Outcome & Impact

This project delivers a prototype system, integrated newly developed algorithms, which is supposed to be commercialised while remaining open to the academia for non-commercial use. Key results are expected to lead to 1-2 patents and 3-5 papers in top venues in machine learning (e.g., ICML, NeurIPS, ICLR), operations research and management science (e.g., *Operations Research* and *Management Science*), and biomedical and health care informatics (e.g., the *New England Journal of Medicine*). The deliverables are set to make significant impacts on the industry and public health care sectors in the UK and the wider world.