International Journal of Information Technology Vol. 29 No. 2 2023

Tailoring ChatGPT for Enhanced Pain Management: A Proof-of-concept Study

Xu Guo, Haocong Rao and Shuaicheng Niu

Nanyang Technological University, Singapore

Abstract

Pain management expertise is mainly captured in clinical notes, which are unstructured and predominantly composed of natural languages written by physicians. Identifying the relationships between patient data and individualized treatment strategies is challenging as it requires a deep understanding of clinical notes. This paper investigates the application of ChatGPT, a popular Large Language Model (LLM) known for its impressive generative abilities in AI, for summarizing pain management data and providing treatment suggestions. Using the templates created by ChatGPT, we successfully extracted paired pain-treatment examples from existing electronic health records. Through in-context learning on the few-shot demonstrations, ChatGPT can generate new personalized treatment plans for new patients.

Keywords: Pain Management, ChatGPT, Generative AI, Large Language Models.

I. Introduction

Pain management is a critical and intricate field in healthcare. Its primary objective is to cater to the diverse and individualized needs of patients experiencing pain [11, 12]. The effective treatment of pain hinges on in-depth understanding of each patient's unique circumstances, including their medical history, pain characteristics, and responses to previous treatments [2, 7]. Such crucial information is often contained in clinical notes, which serve as a rich repository of patient data. These notes, however, are often unstructured and can be overwhelming in quantity, making their analysis a significant challenge in pain management [15, 6].

Over the past year, various LLMs such as LlaMa [18], PaLM[3], BLIP[10] and ChatGPT[13], have brought significant impact on industries. They have showcased impressive capabilities across a wide range of applications, ranging from their proficiency in general content creation[1, 5, 4] to their excellence in delivering personalized solutions in customer service[17, 16] and even personality assessment [14, 9]. Impressed by its strong natural language understanding and generation capabilities, this paper aims to design prompting approaches to tailor ChatGPT to pain management practice. Specifically, we ask ChatGPT to create templates to summarize and reformat existing anonymized electronic health records (EHRs) into pain conditions and treatments pairs. We tailor ChatGPT to the task of pain treatment recommendation [8] by conducting in-context learning on few-shot pain-treatment pairs. In the testing scenarios, only the pain conditions are provided and ChatGPT is required to generate treatment recommendations to assist pain doctors.

II. Methodology

A. Overview

In this paper, we propose a novel approach that conducts in-context learning for GPT-4V using few-shot data from patients experiencing pain. This method improves the ability of GPT-4V to help diagnose, manage, and treat pain, providing useful suggestions in patient care. As shown in Fig. 1, the overall pipeline is: 1) identify and extract key attributes directly linked to pain diagnosis and treatment; 2) summarize the patients' details using templates generated by GPT-4V; 3) construct paired data comprising patient history and their corresponding pain management/treatment plans,

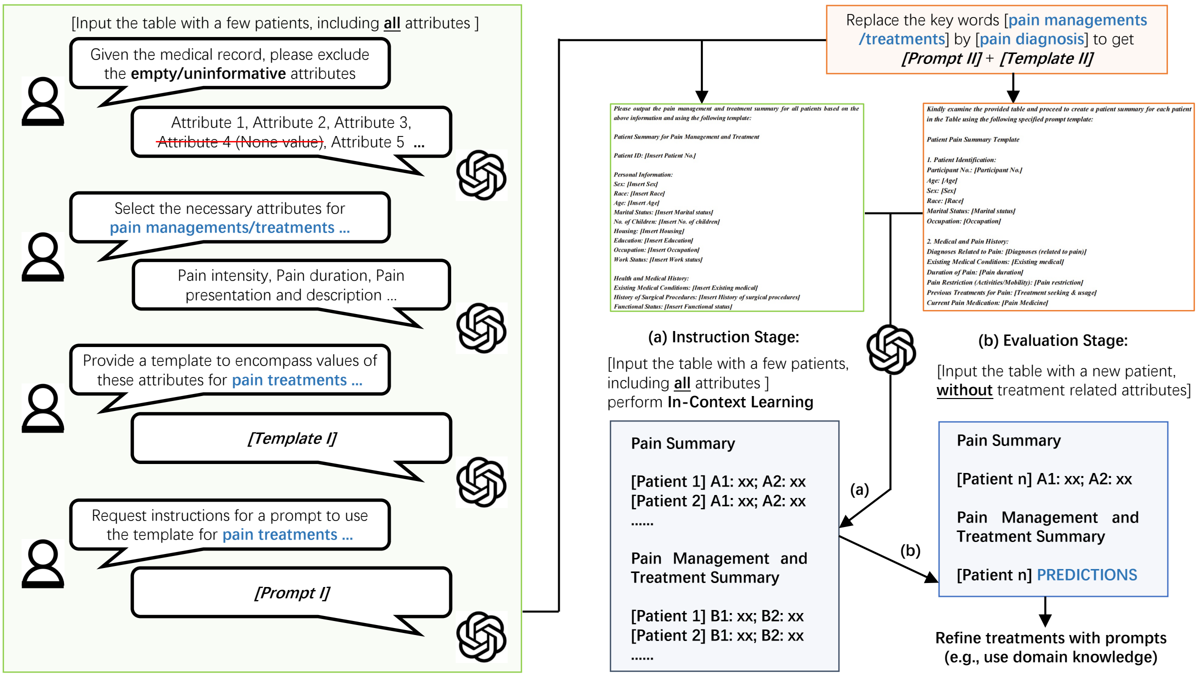

Fig 1 Schematic diagram of the proposed approach for pain treatment.

for the in-context learning of GPT-4V; 4) evaluate the learned GPT-4V on new unseen patient cases. We depict this pipeline in the subsequent sections.

B. Essential Attributes Extraction

The original data, comprising various patients' medical records, often contains numerous missing values and attributes that may not be relevant to pain management and treatment. For convenient understanding and analysis of GPT-4V, we initially utilize GPT-4V to autonomously identify and extract the essential attributes. We format the raw data into a tabular structure and then feed it into GPT-4V. For the extraction of attributes, we employ the following prompt:

"Please read the table and exclude the first-row attributes for which some samples do not possess specific values or text (e.g., not providing any information about the attribute). Then, please output each of the excluded attributes with the detailed reason for exclusion. Finally, please output all included attributes. Each attribute should be output line by line. "

Subsequently, we ask GPT-4V to refine the selection of essential attributes. We use the specific prompt: *"Please select the essential attributes from above that are determined or highly related to the final pain diagnosis/managements and treatments"*. This step ensures a more focused and relevant extraction of attributes.

C. Patients Case Summarization

Utilizing the attributes extracted above, we direct GPT-4V to create templates for summarizing patients' pain diagnoses and pain treatment. Each template is designed to include the patient's basic information, alongside their specific pain diagnosis or treatment details. These structured summaries are intended to 1) aid doctors in more effectively evaluating patient conditions, diagnosing pain, and deciding on suitable treatments; 2) help GPT-4V better understand the patient's medical records. The specific prompts designed for this task are as follows:

"I kindly request a template for summarizing patient details, encompassing the above aspects or attributes. This template should include placeholders that can be filled in with the specific details of each patient. The intention is for these tailored summaries to assist doctors in evaluating the patient's condition and diagnosing pain [or determining appropriate treatments]. Please make sure all information in the template can be extracted from table data."

D. Few-Shot In-Context Learning

For each patient's data, we have developed two distinct templates focusing on pain diagnosis and treatment. The contents in these templates are structured as pairs, encompassing [patient's basic information, corresponding pain diagnosis] and [patient's basic information, corresponding pain treatment]. We input patients' information extracted according to these two templates into GPT-

4V for in-context learning. This process is designed to enable GPT-4V to acquire specialized knowledge in the domain of pain diagnosis and treatment. To facilitate this learning, we utilize the following prompts:

"Kindly examine the provided table and proceed to create a patient diagnosis summary for each patient in the Table using the following specified prompt template [generated above]" and *"Please output the pain management and treatment summary for all patients based on the above information and using the following template [generated above]"*.

Next, we input the patients' information (structured by the above templates) into GPT-4V, tasking it to comprehend the knowledge of pain diagnosis and treatment. To guide this process, we employ the following prompt: *"Kindly take note of the above patients' details, including how to diagnose and treat pain. Following this, I will present information about another patient, and I would appreciate your assistance in proposing corresponding pain treatment recommendations."*

E. Evaluation

For the evaluation phase, we input the patient's information into GPT-4V, formatted the same as the previously used tables, but excluding treatment contents. Our objective is to recommend pain treatments. We accomplish this by first extracting patient information using the same pain treatment template created earlier. Then, we prompt GPT-4V to fill in the attributes marked as "null/not indicated." The specific prompts used for this purpose are detailed as:

"Please output the pain management and treatment summary for two patients using the above Patient Summary for Pain Management and Treatment template. Note that some attributes related to treatments are masked, please use the above patient information, context information, and your domain knowledge to generate."

Note that the generated pain treatments can be further refined by adding prompts ($e.g.$, use context information and domain knowledge).

III. Experiments

A. Datasets

We conduct experiments on a self-collected pain management dataset that has a total of 5 participants. The dataset provides a detailed compilation of demographic, personal, and health-related information on various participants. It includes attributes such as participant sex, race, age, marital status, number of children, housing situation, education level, and occupation. Additionally, it encompasses detailed medical consultation histories, including the number of specialists consulted, as well as specific reviews in orthopaedics, rheumatology, neurology, cardiology, physiotherapy, psychology, occupational therapy, and podiatry.

B. Results

We utilize the data from three participants in the dataset as training data for in-context learning with GPT-4V, while the data from two other participants, stripped of their pain treatment details, are used for testing. Our approach begins by identifying key attributes closely related to pain diagnosis and treatment using GPT-4V, as shown in the patient summary template in Fig. 2. Subsequently, we construct a pain diagnosis and treatment template to conduct in-context learning. After that, we present GPT-4V with the task of filling out this template for new unseen patients, thereby generating appropriate treatment recommendations (detailed results are depicted in Fig. 2). The templates and treatment plans generated for these unseen patients underscore the effectiveness of

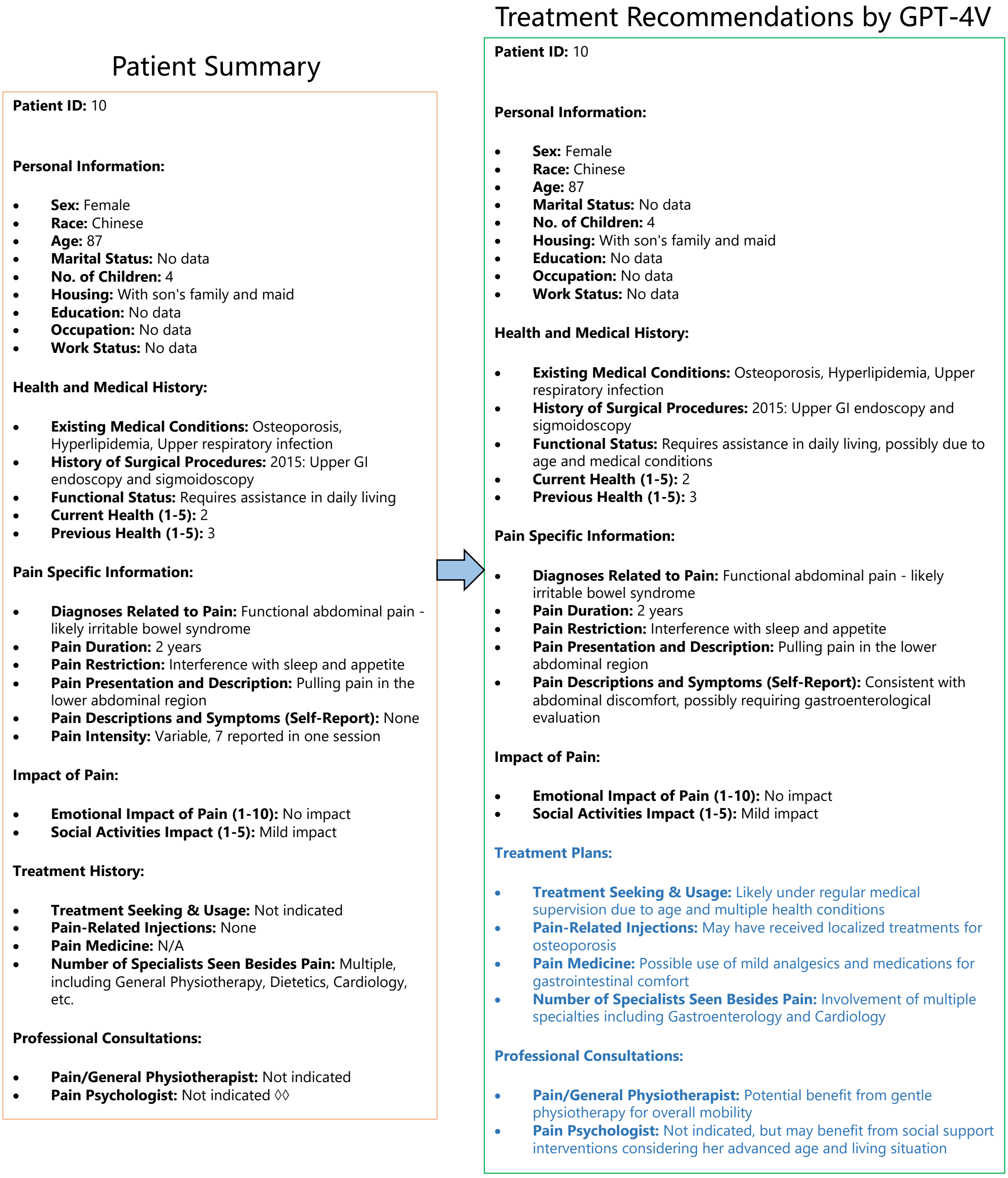

Fig 2 An illustration of GPT-4V generated treatment recommendations for an unseen patient.

targeted prompting techniques in enabling GPT-4V to provide well-reasoned suggestions in the realm of pain management and treatment.

IV. Conclusion and Future Work

This paper provides a proof-of-concept study on a few-shot anonymized dataset for the application of ChatGPT (GPT4V) in pain management. We develop a pipeline that extracts structured data from unstructured clinical notes as few-shot demonstrations for ChatGPT to generate personalized treatment strategies through in-context learning. Empirical experiments reveal that ChatGPT can identify key pain attributes and provide detailed treatment plans. Our study demonstrates its potential to assist pain doctors and improve efficiency in pain management. In the future work, we will explore integrating chain-of-thought reasoning to enhance the coherence between the AI-generated treatment strategies and doctor's decision-making processes.

References

[1] Yihan Cao, Siyu Li, Yixin Liu, Zhiling Yan, Yutong Dai, Philip S Yu, and Lichao Sun. A comprehensive survey of ai-generated content (aigc): A history of generative ai from gan to chatgpt. *arXiv preprint arXiv:2303.04226*, 2023.

[2] Eloise Carr. Barriers to effective pain management. *Journal of Perioperative Practice*, 17(5):200–208, 2007.

[3] Aakanksha Chowdhery, Sharan Narang, Jacob Devlin, Maarten Bosma, Gaurav Mishra, Adam Roberts, Paul Barham, Hyung Won Chung, Charles Sutton, Sebastian Gehrmann, Parker Schuh, Kensen Shi, Sasha Tsvyashchenko, Joshua Maynez, Abhishek Rao, Parker Barnes, Yi Tay, Noam Shazeer, Vinodkumar Prabhakaran, Emily Reif, Nan Du, Ben Hutchinson, Reiner Pope, James Bradbury, Jacob Austin, Michael Isard, Guy Gur-Ari, Pengcheng Yin, Toju Duke, Anselm Levskaya, Sanjay Ghemawat, Sunipa Dev, Henryk Michalewski, Xavier Garcia, Vedant Misra, Kevin Robinson, Liam Fedus, Denny Zhou, Daphne Ippolito,

David Luan, Hyeontaek Lim, Barret Zoph, Alexander Spiridonov, Ryan Sepassi, David Dohan, Shivani Agrawal, Mark Omernick, Andrew M. Dai, Thanumalayan Sankaranarayana Pillai, Marie Pellat, Aitor Lewkowycz, Erica Moreira, Rewon Child, Oleksandr Polozov, Katherine Lee, Zongwei Zhou, Xuezhi Wang, Brennan Saeta, Mark Diaz, Orhan Firat, Michele Catasta, Jason Wei, Kathy Meier-Hellstern, Douglas Eck, Jeff Dean, Slav Petrov, and Noah Fiedel. Palm: Scaling language modeling with pathways, 2022.

[4] Zilin Du, Haoxin Li, Xu Guo, and Boyang Li. Training on synthetic data beats real data in multimodal relation extraction. *arXiv preprint arXiv:2312.03025*, 2023.

[5] Zilin Du, Yunxin Li, Xu Guo, Yidan Sun, and Boyang Li. Training multimedia event extraction with generated images and captions. *arXiv preprint arXiv:2306.08966*, 2023.

[6] Samah Jamal Fodeh, Dezon Finch, Lina Bouayad, Stephen L Luther, Han Ling, Robert D Kerns, and Cynthia Brandt. Classifying clinical notes with pain assessment using machine learning. *Medical & biological engineering & computing*, 56:1285–1292, 2018.

[7] Diane Glowacki. Effective pain management and improvements in patients' outcomes and satisfaction. *Critical care nurse*, 35(3):33–41, 2015.

[8] Xu Guo, Han Yu, Chunyan Miao, and Yiqiang Chen. Agent-based decision support for pain management in primary care settings. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19*, pages 6521–6523. International Joint Conferences on Artificial Intelligence Organization, 7 2019.

[9] Guangyuan Jiang, Manjie Xu, Song-Chun Zhu, Wenjuan Han, Chi Zhang, and Yixin Zhu. Mpi: Evaluating and inducing personality in pre-trained language models. *arXiv preprint arXiv:2206.07550*, 2022.

[10] Junnan Li, Dongxu Li, Caiming Xiong, and Steven Hoi. Blip: Bootstrapping language-image

pre-training for unified vision-language understanding and generation, 2022.

[11] Laxmaiah Manchikanti, Vanila Singh, Alan D Kaye, and Joshua A Hirsch. Lessons for better pain management in the future: learning from the past. *Pain and Therapy*, 9:373–391, 2020.

[12] Edward R Mariano, David M Dickerson, Joseph W Szokol, Michael Harned, Jeffrey T Mueller, Beverly K Philip, Jaime L Baratta, Padma Gulur, Jennifer Robles, Kristopher M Schroeder, et al. A multisociety organizational consensus process to define guiding principles for acute perioperative pain management. *Regional Anesthesia & Pain Medicine*, 2021.

[13] OpenAI. Introducing chatgpt, 2023.

[14] Haocong Rao, Cyril Leung, and Chunyan Miao. Can ChatGPT assess human personalities? a general evaluation framework. In Houda Bouamor, Juan Pino, and Kalika Bali, editors, *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 1184–1194, Singapore, December 2023. Association for Computational Linguistics.

[15] Joanne G Samuels and Dawn Kritter. Pain management documentation: analyzing one hospital's computerized clinical records. *CIN: Computers, Informatics, Nursing*, 29(9):512–518, 2011.

[16] Agus Dedi Subagja, Abu Muna Almaududi Ausat, Ade Risna Sari, M Indre Wanof, and Suherlan Suherlan. Improving customer service quality in msmes through the use of chatgpt. *Jurnal Minfo Polgan*, 12(2):380–386, 2023.

[17] Frans Sudirjo, Karno Diantoro, Jassim Ahmad Al-Gasawneh, Hizbul Khootimah Azzaakiyyah, and Abu Muna Almaududi Ausat. Application of chatgpt in improving customer sentiment analysis for businesses. *Jurnal Teknologi Dan Sistem Informasi Bisnis*, 5(3):283–288, 2023.

[18] Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Tim-

othée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, Aurelien Rodriguez, Armand Joulin, Edouard Grave, and Guillaume Lample. Llama: Open and efficient foundation language models, 2023.