Title: Using Large Language Models (LLMs) to Detect Therapeutic Inertia in Hypertension Management

Keywords: Natural Language Processing (NLP), Large Language Models (LLMs), Clinical Natural Language Processing (Clinical NLP), Therapeutic Inertia, Hypertension

# Abstract

Therapeutic inertia is the failure to intensify antihypertensive treatment despite uncontrolled blood pressure. This can lead to increased cardiovascular risk and mortality. Large language models (LLMs) have the potential to be used to detect therapeutic inertia by extracting and analyzing clinical information from progress notes.

This paper proposes a method for using LLMs to detect therapeutic inertia in hypertension management. The method first preprocesses the progress notes to extract clinical information. The clinical information is then used to train an LLM to identify topics related to therapeutic inertia. The method was evaluated on a dataset of progress notes from patients with hypertension.

The data was pre-processed using regular expressions and the NLP models. Topic modeling was then performed using transformer models such as BERTopic and c-TF-IDF to construct dense clusters and analyze top topics. The results showed that Topic .., with the terms …., had the best coherence score of …. In contrast, Topic …, with the terms …. Received the lowest coherence score of …. The proposed method has the potential to be used to improve the management of hypertension by identifying patients who are at risk of therapeutic inertia.

# Introduction

Large language models (LLM) have recently made significant strides in improving the performance of a range of natural language processing (NLP) tasks, revealing new possibilities for the automation of tasks that were previously completed by humans. LLMs can significantly change the healthcare sector by gleaning insightful information from unstructured data, including digital medical records and electronic health records. Through the identification of critical data points for clinical trials, population health management, and drug discovery, LLMs may help in the development of novel medications and treatment strategies. Professionals can anticipate notable advancements in patient outcomes and overall healthcare delivery if LLMs are developed and integrated into the healthcare system.

When it comes to emergent skills in zero-shot learning, LLMs have a clear advantage because they can learn and adjust to new tasks with prompt instructions, even if they have never seen them before [4, 5]. For instance, ChatGPT can effectively compete with commercial translation tools like Google Translate by integrating the instruction prompt "Translate these sentences from [source language] to [target language]" [6]. Although LLM is a useful tool for users because it can perform a wide range of NLP tasks, it also raises serious privacy concerns. A significant worry is that during the procedure, private information might unintentionally come to light. This is especially true in the healthcare industry, where maintaining patient privacy and confidentiality is crucial. Consequently, it's critical to make sure that it is there.

In order to assess the zero-shot performance of current LLM models for healthcare tasks, we conducted experiments on LLMs to investigate its ability to extract structured information from unstructured healthcare texts, specifically for Named Entity Recognition (NER) tasks. Our preliminary findings suggest that using topic modeling directly only yields poor performance compared to SOTA models trained on the progress note. This result highlights the fact that while has demonstrated impressive inference and reasoning

abilities in various classic Natural Language Understanding (NLU) tasks, it is not adequate to apply berTopic alone to healthcare tasks since it was not specifically trained for this domain [7, 8]. As a result, it's critical to make sure that there is a system in place to guarantee strong privacy protections that stop illegal access to sensitive data. When utilizing LLM, reliability and usability are also crucial concerns that need to be taken into consideration. The model's accuracy and consistency must be relied upon by users, which calls for constant testing, evaluation, and improvement to make sure the system is operating as planned and serving users' needs. In this work, we will concentrate on enhancing LLM's dependability for NER tasks when we just exclude individual names..

In conclusion, our innovative training framework enables the training of a local model with superior performance compared to using LLMs alone. It also mitigates potential privacy concerns and reduces the dependence on costly and time-consuming data collection and labeling.

# Preliminaries

**Biomedical Named Entity Recognition.** Biomedical NER involves identifying and categorizing medical entities, such as diseases, symptoms, drugs, etc., in a medical text. NER uses an IOB (Inside, Outside, Begin) tagging scheme, where each word is assigned a tag indicating whether it is the beginning of a named entity (B), inside a named entity

**Biomedical Relation Extraction.** Biomedical RE involves identifying and extracting the relationships between medi- cal entities in a text, such as diseases and drugs, symptoms and treatments, etc. Formally, let *x* be a sentence containing.

**Zero-shot Learning.** Zero-shot learning is an emerging research paradigm that allows LLMs to perform tasks they have not been explicitly trained. This is accomplished by utilizing the LLMs’ capacity to produce coherent text based on a given prompt. The prompt serves as a guide, providing a corpus that describes the task at hand along with a set of potential outputs [5]. The LLM then generates the most plausible output based on its acquired knowledge. Recent studies show that LLMs achieve promising zero-shot ability in various traditional NLU tasks [10].

# Benchmarking LLM on Biomedical NER

We conducted benchmark experiments on the ChatGPT model. To create prompts that would effectively triggers named entity recognition.

A white rectangular object with black text

Description automatically generated

A diagram of a model

Description automatically generated

# Named Entity Recognition

In this section, we evaluate the effectiveness of our proposed synthetic data generation approach for the named entity recognition (NER) task, following the methodology outlined in Section 4. We first extract the *M* seed entities from the training set and use them to generate synthetic sentences with annotations for the target entity type. Specifically,

set *N* = 30. We then use the synthetic dataset to fine-tune three pre-trained language models. To evaluate the for

each seed entity, we generate *N* sentences with the corresponding entity annotations, in our experiments, we

performance of the baseline models and our proposed methods, we use a subset of the test set for all datasets due

to the computational limitations of ChatGPT. We report the precision, recall, and F1 scores for each model. The

models evaluated in our experiments include three settings: (1) zero-shot, where the models were not fine-tuned on

any dataset, and the model is ChatGPT. (2) models fine-tuned on synthetic data generated by our approach, and (3)

models fine-tuned on the original training set. The pre-trained language models used in our experiments are BERT

[18], RoBERTa [19], and BioBERT [20].

**NLP for Biomedical.** The Natural Language Processing (NLP) technique is widely applied in the biomedical domain, as evidenced by numerous studies [28, 29]. NLP for Biomedical has various applications, including the analysis of electronic health records (EHRs)[30, 31, 32, 33], drug discovery[34, 35], and medical chatbots [36, 37]. The use of LLMs for Biomedical is gaining traction among both industry and academic researchers. Previous work [20, 38, 39] has explored the application of NER and RE to biomedical tasks. Biomedical NER and RE has diverse usage in the healthcare domain, including analyzing EHRs [40, 41, 42], extracting clinical trials [43, 44], and drug development [45, 46, 47]. The major challenges facing NLP in Biomedical include developing accurate models for biomedical text analysis and ensuring patient data privacy. In this work, we propose the use of synthetic data to fine-tune offline models, which can not only improve prediction accuracy but also protect patient privacy.

# References

* 1. Rosenblatt K. ChatGPT passes MBA exam given by a Wharton professor; 2023. Available from: [https://www](http://www/). nbcnews.com/tech/tech-news/chatgpt-passes-mba-exam-wharton-professor-rcna67036.
  2. Choi JH, Hickman KE, Monahan A, Schwarcz D. ChatGPT Goes to Law School. Available at SSRN. 2023. [3]Baidoo-Anu D, Owusu Ansah L. Education in the Era of Generative Artificial Intelligence (AI): Understanding

the Potential Benefits of ChatGPT in Promoting Teaching and Learning. Available at SSRN 4337484. 2023. [4]Wei J, Tay Y, Bommasani R, Raffel C, Zoph B, Borgeaud S, et al. Emergent abilities of large language models.

arXiv preprint arXiv:220607682. 2022.

1. Liu P, Yuan W, Fu J, Jiang Z, Hayashi H, Neubig G. Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. ACM Computing Surveys. 2023;55(9):1-35.
2. Jiao W, Wang W, Huang Jt, Wang X, Tu Z. Is ChatGPT a good translator? A preliminary study. arXiv preprint arXiv:230108745. 2023.
3. Brown T, Mann B, Ryder N, Subbiah M, Kaplan JD, Dhariwal P, et al. Language models are few-shot learners.

Advances in neural information processing systems. 2020;33:1877-901.

1. Qin C, Zhang A, Zhang Z, Chen J, Yasunaga M, Yang D. Is ChatGPT a General-Purpose Natural Language Processing Task Solver? arXiv preprint arXiv:230206476. 2023.
2. Aziz HA. A review of the role of public health informatics in healthcare. Journal of Taibah University Medical Sciences. 2017;12(1):78-81.
3. Zhong Q, Ding L, Liu J, Du B, Tao D. Can ChatGPT Understand Too? A Comparative Study on ChatGPT and Fine-tuned BERT. arXiv preprint arXiv:230210198. 2023.
4. Li J, Sun Y, Johnson RJ, Sciaky D, Wei CH, Leaman R, et al. BioCreative V CDR task corpus: a resource for chemical disease relation extraction. Database. 2016;2016.
5. Do g˘an RI, Leaman R, Lu Z. NCBI disease corpus: a resource for disease name recognition and concept normal- ization. Journal of biomedical informatics. 2014;47:1-10.
6. Rouillard AD, Gundersen GW, Fernandez NF, Wang Z, Monteiro CD, McDermott MG, et al. The harmonizome: a collection of processed datasets gathered to serve and mine knowledge about genes and proteins. Database. 2016;2016.
7. Van Mulligen EM, Fourrier-Reglat A, Gurwitz D, Molokhia M, Nieto A, Trifiro G, et al. The EU-ADR corpus: annotated drugs, diseases, targets, and their relationships. Journal of biomedical informatics. 2012;45(5):879-84.
8. Guo B, Zhang X, Wang Z, Jiang M, Nie J, Ding Y, et al. How Close is ChatGPT to Human Experts? Comparison Corpus, Evaluation, and Detection. arXiv preprint arXiv:230107597. 2023.
9. Voigt P, Von dem Bussche A. The eu general data protection regulation (gdpr). A Practical Guide, 1st Ed, Cham: Springer International Publishing. 2017;10(3152676):10-5555.
10. Goldman E. An introduction to the california consumer privacy act (CCPA). Santa Clara Univ Legal Studies Research Paper. 2020.
11. Devlin J, Chang MW, Lee K, Toutanova K. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:181004805. 2018.
12. Liu Y, Ott M, Goyal N, Du J, Joshi M, Chen D, et al. Roberta: A robustly optimized bert pretraining approach. arXiv preprint arXiv:190711692. 2019.
13. Lee J, Yoon W, Kim S, Kim D, Kim S, So CH, et al. BioBERT: a pre-trained biomedical language representation model for biomedical text mining. Bioinformatics. 2020;36(4):1234-40.
14. Kasneci E, Seßler K, K u¨chemann S, Bannert M, Dementieva D, Fischer F, et al. ChatGPT for good? On opportunities and challenges of large language models for education. EdArXiv. 2023.
15. Arora A, Arora A. The promise of large language models in health care. The Lancet. 2023;401(10377):641. [23]Singhal K, Azizi S, Tu T, Mahdavi SS, Wei J, Chung HW, et al. Large Language Models Encode Clinical

Knowledge. arXiv preprint arXiv:221213138. 2022.

1. Yuan A, Coenen A, Reif E, Ippolito D. Wordcraft: story writing with large language models. In: 27th Interna- tional Conference on Intelligent User Interfaces; 2022. p. 841-52.
2. Scao TL, Fan A, Akiki C, Pavlick E, Ili c´ S, Hesslow D, et al. Bloom: A 176b-parameter open-access multilingual language model. arXiv preprint arXiv:221105100. 2022.
3. Chowdhery A, Narang S, Devlin J, Bosma M, Mishra G, Roberts A, et al. PaLM: Scaling Language Modeling

with Pathways. CoRR. 2022;abs/2204.02311. Available from: https://doi.org/10.48550/arXiv.2204.02311. [27]Touvron H, Lavril T, Izacard G, Martinet X, Lachaux MA, Lacroix T, et al. LLaMA: Open and Efficient Foun-

dation Language Models. arXiv preprint arXiv:230213971. 2023.

[28]Yandell MD, Majoros WH. Genomics and natural language processing. Nature Reviews Genetics.

2002;3(8):601-10.

[29]N e´dellec C, Bossy R, Kim JD, Kim JJ, Ohta T, Pyysalo S, et al. Overview of BioNLP shared task 2013. In: Proceedings of the BioNLP shared task 2013 workshop; 2013. p. 1-7.

1. Koleck TA, Dreisbach C, Bourne PE, Bakken S. Natural language processing of symptoms documented in free- text narratives of electronic health records: a systematic review. Journal of the American Medical Informatics Association. 2019;26(4):364-79.
2. Patra BG, Sharma MM, Vekaria V, Adekkanattu P, Patterson OV, Glicksberg B, et al. Extracting social determi- nants of health from electronic health records using natural language processing: a systematic review. Journal of the American Medical Informatics Association. 2021;28(12):2716-27.
3. Ohno-Machado L. Realizing the full potential of electronic health records: the role of natural language process- ing. Journal of the American Medical Informatics Association. 2011;18(5):539-9.
4. Wang X, Hripcsak G, Markatou M, Friedman C. Active computerized pharmacovigilance using natural lan- guage processing, statistics, and electronic health records: a feasibility study. Journal of the American Medical Informatics Association. 2009;16(3):328-37.
5. O¨ ztu¨rk H, O¨ zgu¨r A, Schwaller P, Laino T, Ozkirimli E. Exploring chemical space using natural language pro-

cessing methodologies for drug discovery. Drug Discovery Today. 2020;25(4):689-705.

1. Vamathevan J, Clark D, Czodrowski P, Dunham I, Ferran E, Lee G, et al. Applications of machine learning in drug discovery and development. Nature reviews Drug discovery. 2019;18(6):463-77.
2. Safi Z, Abd-Alrazaq A, Khalifa M, Househ M, et al. Technical aspects of developing chatbots for medical applications: scoping review. Journal of medical Internet research. 2020;22(12):e19127.
3. Dharwadkar R, Deshpande NA. A medical chatbot. International Journal of Computer Trends and Technology (IJCTT). 2018;60(1):41-5.
4. Yu X, Hu W, Lu S, Sun X, Yuan Z. BioBERT based named entity recognition in electronic medical record. In: 2019 10th international conference on information technology in medicine and education (ITME). IEEE; 2019. p. 49-52.
5. Zhu Y, Li L, Lu H, Zhou A, Qin X. Extracting drug-drug interactions from texts with BioBERT and multiple entity-aware attentions. Journal of biomedical informatics. 2020;106:103451.
6. Gorinski PJ, Wu H, Grover C, Tobin R, Talbot C, Whalley H, et al. Named entity recognition for electronic health records: a comparison of rule-based and machine learning approaches. arXiv preprint arXiv:190303985. 2019.
7. Dai Z, Wang X, Ni P, Li Y, Li G, Bai X. Named entity recognition using BERT BiLSTM CRF for Chinese electronic health records. In: 2019 12th international congress on image and signal processing, biomedical engineering and informatics (cisp-bmei). IEEE; 2019. p. 1-5.
8. Kormilitzin A, Vaci N, Liu Q, Nevado-Holgado A. Med7: A transferable clinical natural language processing model for electronic health records. Artificial Intelligence in Medicine. 2021;118:102086.
9. Kang T, Zhang S, Tang Y, Hruby GW, Rusanov A, Elhadad N, et al. EliIE: An open-source information ex- traction system for clinical trial eligibility criteria. Journal of the American Medical Informatics Association. 2017;24(6):1062-71.
10. Nye BE, DeYoung J, Lehman E, Nenkova A, Marshall IJ, Wallace BC. Understanding clinical trial re- ports: Extracting medical entities and their relations. AMIA Summits on Translational Science Proceedings. 2021;2021:485.
11. Shinozaki A. Electronic medical records and machine learning in approaches to drug development. In: Artificial intelligence in Oncology drug discovery and development. IntechOpen; 2020. .
12. Rockt a¨schel T, Weidlich M, Leser U. ChemSpot: a hybrid system for chemical named entity recognition. Bioin- formatics. 2012;28(12):1633-40.
13. Eltyeb S, Salim N. Chemical named entities recognition: a review on approaches and applications. Journal of cheminformatics. 2014;6:1-12.