medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308731; this version posted June 11, 2024. The copyright holder has placed this preprint (which was not certified by peer review) in the Public Domain. It is no longer restricted by copyright. Anyone can legally share, reuse, remix, or adapt this material for any purpose without crediting the original authors.

Impact of Language Models on Healthcare in Thailand: Benefits, Challenges, and Future Opportunities

Preprint, compiled June 11, 2024

Chanakan Moongtin [1*]

[1]Faculty of Information and Communication Technology, Mahidol University, Nakhon Pathom, Thailand

Abstract

This study explores the use of Artificial Intelligence (AI), specifically Large Language Models (LLMs) in the Thai healthcare sector, focusing on applications such as diagnosis, patient monitoring, and automated question-and-answer systems. While AI has the potential to improve diagnosis accuracy, reduce the time required for appointment, and enhance patient care, several challenges prevent widespread adoption of LLMs in healthcare, including significant computational resources required for deployment, data privacy and security concerns, and Thai language being a low-resource language. Through a comprehensive analysis of publicly available online data and literature, this study examines the current state of AI adoption in Thai healthcare, identifying key barriers to adoption and providing recommendations for overcoming these challenges, including targeted training and education for healthcare professionals, strategic government initiatives, and investments in infrastructure. By addressing these issues, Thailand can harness the full potential of AI technologies to enhance its healthcare system, ensuring better patient outcomes and operational efficiencies.

Keywords Large language models, LLMs, healthcare, Thailand, AI, diagnosis, ML, pathology

1 Introduction

The rapid advancement of Artificial Intelligence (AI) has transformed various industries, and healthcare is no exception. From Recurrent Neural Network to Transformer [1], LLMs can understand and respond to human natural languages with an accuracy higher than before. Sparking the potential for usage in applications such as chatbots, document summarization, and risk analysis. The integration of AI in healthcare has the potential to revolutionize the sector, improving diagnosis accuracy, enhancing patient care, and streamlining clinical workflows. Certain LLMs that have not been trained on medical data, such as GPT-4, have nevertheless been successfully utilized by medical professionals and researchers to analyze and interpret medical data, revealing their potential utility in both research and practical applications [2].

However, the adoption of AI in healthcare is not without its challenges, particularly in developing countries like Thailand. This study aims to explore the current state of AI adoption in the Thai healthcare sector, identifying the key barriers to adoption and providing recommendations for overcoming these challenges.

2 Opportunities and Challenges of using LLMs within healthcare in Thailand

Large Language Models currently have its pros and cons to consider when using these models in certain applications. The potentials of using LLMs in healthcare field are as follows [3]:

2.1 *Opportunities*

1. Vast Information and Knowledge: LLMs have been trained on large corpus of internet data, which include medical literature, research, papers and other relevant sources, allowing LLMs to respond and analyze certain context related to medical field.
2. Accessibility and Availability: Language models are different to human, they are mostly always online and respond at speed higher than any human can, and it is also much easier and more accessible to access LLMs than to arrange an appointment with the doctors where a lot of factors could make or break an appointment.
3. LLMs can keep learning new things: LLMs can be trained and updated with the latest medical data, ensuring that LLMs always have the latest information needed, allowing for flexibility and adaptability to improve the model over time.

Despite the potential to revolutionize healthcare for everyday usage, and the applications that LLMs can be useful for, several obstacles prevent the widespread adoption of LLMs within Thailand, such as:

2.2 *Challenges*

1. Ethical Concerns regarding Privacy and Security: Since LLMs are usually large (as the name suggest), and thus require a lot of resources to run, it is usually not feasible to deploy LLMs on-premise due to cost and power usage, which encourage people to deploy them on the cloud. This approach resulted in privacy and security concerns of relying on 3rd party to process users data, especially sensitive information such as, but not limited to, user's diagnosis information, user's past diagnosis, etc. Making the usage of 3rd party platforms for processing and storing sensitive data impractical. And there can be concerns related to

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308731; this version posted June 11, 2024. The copyright holder has placed this preprint (which was not certified by peer review) in the Public Domain. It is no longer restricted by copyright. Anyone can legally share, reuse, remix, or adapt this material for any purpose without crediting the original authors.

potential liability and accountability of AI systems in healthcare that have to be addressed in order to avoid potential harm to patients and legal implications.

2. Accuracy Problems: LLMs are usually infamous for its propensity to hallucinate. Hallucinations, as defined by Berrios and Dening [4] quoted "Hallucinations are conceived of as indistinguishable from real perceptions except that there is no stimulus", and in the case of LLMs, it means that the model will produce inaccurate or nonsensical information regarding the topics. Which can be caused by low amount of information related to the topic, bias within the base model itself such as memorization at the level of sentences, and statistical patterns of usage learned at the level of corpora [5].
3. Thai is a low-resource language: When using Language Models in Thailand, having the ability to understand Thai language and context is important. Unlike English, Thai has received a lot less attention from the NLP community. Which makes finding large corpus of Thai contents to pre-train LLMs on, a more complicated task. As such, there weren't any Thai language models that fully captures Thai context and culture. But recently, there have been progress on creating Thai large language model from multilingual datasets such as MC4, which is the cleaned version of Common Crawl's web crawl corpus, and Oscar, which is an open-source multilingual dataset. The work to create a better-understanding LLM for Thai culture is continuously being done under the name Typhoon [6] by SCB 10X, in which we hope can help mitigate these problems.

3 EXISTING SOLUTIONS

While several challenges hinder the adoption of Large Language Models (LLMs) in Thai healthcare, some concerns can be mitigated through innovative solutions. For instance, at the basic, giving patients a warning about the service utilizing LLMs to process patients data, giving patients a choice to consent, then privacy and security concerns can be alleviated by employing smaller language models with moderate parameter counts, and enabling the model to reference external information while responding to users. One viable option is to utilize a Thai language model, such as LLaMA-3-Typhoon-v1.5-8B or LLaMA-3-Typhoon-v1.5-8B-Instruct [6], and fine-tune it for the goal of enabling LLM to respond within the desired patterns, and not necessary for imparting knowledge, as any knowledge imparted into fine-tuning steps will result in knowledge degradation to the model [7]. Subsequently, quantization techniques such as GGUF from llama.cpp [8] can be applied to reduce the model size, allowing it to be run on lower-resource system while maintaining its level of accuracy. To further mitigate accuracy problems, during inference, a retrieval mechanism can be added to the LLM, allowing it to recall data from vector storage solutions such as ChromaDB [9] or Weaviate [10]. This approach, known as Retrieval-Augmented Generation (RAG), is preferred over fine-tuning knowledge into the base model, as RAG provides a ground-truth source of information to the LLMs [11].

4 METHODOLOGY

4.1 *Objective and Methods*

The objective of this review is to explore the current state of research on the utilization of Large Language Models in healthcare within Thailand. Specifically, this review will assess the potential benefits, address the challenges, and implications of integrating LLMs into healthcare sector.

To conduct this review, a comprehensive search was conducted using all major electronic databases, including Google Scholar, ThaiJO. The keywords used were

- "Language Models" AND "medical" (AND "Thailand" outside ThaiJO)
- "ChatGPT" OR "GPT" AND "medical" (AND "Thailand" outside ThaiJO)
- "ChatGPT" OR "GPT" AND "medical" (In Thai language)

And retrieve relevant articles published within the past five years. The selected articles were then critically analyzed to extract key findings and insights. The search is conducted on June 9, 2024. The records are then selected using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, to ensure the highest quality and transparency in reporting.

4.2 *Criteria*

The systematic review was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines. We use Google Scholar and ThaiJO to search candidate publications. Filters are used when searching on Google Scholar, to enable searching for articles ordered by relevance and published after 2022. The search is conducted on June 9, 2024. The eligibility criteria involved any type of published scientific research or preprints (article, review, communication, etc.) addressing the use of LLMs that fell under the medical field categories within Thailand.

The inclusion criteria are as follows:

1. Papers must be published or reported between 2020 and 2024
2. The context of the paper is about the utilization of Large Language Model (LLMs) in healthcare within Thailand.
3. The paper must be written and published in Thai or English language.
4. The paper is original and published as journal article.

The exclusion criteria included: records that are not in Thai and English, records addressing the use of LLMs in subjects other than those related to medical field, records addressing the use of LLMs in medical field outside of Thailand, and articles from non-academic sources. We excluded 116 out of 132 screened articles, leaving us with 16 full-text articles to be assessed for eligibility.

After completing the eligibility assessment during the selection process (n=7), we then proceed to review the articles as follows.

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308731; this version posted June 11, 2024. The copyright holder has placed this preprint (which was not certified by peer review) in the Public Domain. It is no longer restricted by copyright. Anyone can legally share, reuse, remix, or adapt this material for any purpose without crediting the original authors.

4.3 Results

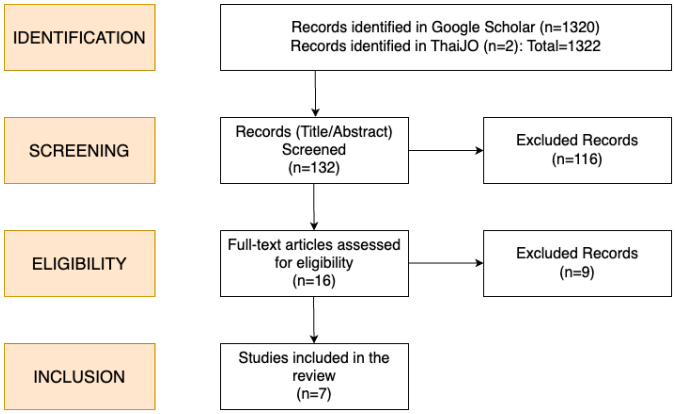

Figure 1: Flowchart of the record selection process based on PRISMA guidelines.

Initially, we have a total of 2 results from ThaiJO, and 1,320 results from Google Scholar (For the first keyword).

5 CRITICAL ANALYSIS OF KEY REVIEWS IN THIS AREA

Table 1 outlines the key reviews published to date that provide examples of the LLMs applications in healthcare within Thailand.

From the analysis, we found below key areas to form a generalized framework:

- Language Model Selection and Fine-tuning: The choice of language model architecture (e.g., XLM-RoBERTa) and fine-tuning on specific tasks (e.g., depression classification) can significantly impact performance, as certain pre-trained language models may not perform well on specific tasks or languages (e.g., Thai) without sufficient training data, and certain models will have different context window, so for some tasks that require processing large amount of data within one context will require language models with larger context window. Certain LLMs will be multimodal, which means the model will accept more than just text (e.g., Visual and/or Auditory) which can be helpful in certain tasks.
- Determine the usage of LLMs in certain tasks: Certain tasks such as depression screening can be done with LLMs, but certain other tasks that requires classification with overlapping symptoms may not be as accurate or useful, it is important to keep every possible parameters that can caused a specific conditions into mind.
- Data Quality and Availability: The availability and quality of training data, especially for low-resource language (e.g., Thai), can significantly impact model performance, as the lack of large, diverse datasets for specific tasks can hinder model development.
- Addressing the Rarity of Diseases and Insufficient Data: As language models work based on the data being provided, having insufficient data can hinder model performance, especially within healthcare and medical field where there may not be enough information for rarer diseases.
- Handling Misinformation and Noise in Data: Models should be designed to handle misinformation, inorganic, or automatically generated data to ensure accurate interpretations.
- Evaluate model performance on multilingual data: Language models should be fine-tuned to the specific language they are going to be used in, to help increase accuracy and better understanding of that language's context.
- Clinical Validation: Collaborate with healthcare professionals to validate the clinical relevance and accuracy of LLMs-based diagnosis tools, to understand their perspectives on integrating such tools into real-world medical practices.
- Ethical Considerations: Address ethical considerations related to patient data privacy, especially Personally Identifiable Information (PII) that LLMs will have access to on certain tasks, and also the bias that will exist in the base model that can cause deviation to the answer.

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308731; this version posted June 11, 2024. The copyright holder has placed this preprint (which was not certified by peer review) in the Public Domain. It is no longer restricted by copyright. Anyone can legally share, reuse, remix, or adapt this material for any purpose without crediting the original authors.

Table 1: Critical Analysis of usage of LLMs application in healthcare in Thailand

Author	Study Title	Area/Focus	Outcomes	Challenges
Suvirat et al. [12]	Leveraging Language Models for Inpatient Diagnosis Coding	The study explores the application of language models in inpatient diagnosis coding using clinical text data that combines both English and Thai words	The study highlights the ability to capture relevant associations, and strong performance with both Thai and mixed English-Thai clinical notes. The study highlights the need for larger datasets to handle rarer diseases, while keeping the document length down as longer documents could pose a challenge to the model.	LLMs have limited context window, in which they cannot handle documents larger than certain length, and the rarity of certain diseases poses a challenge for accurate coding due to lack of information in the training data.
Munthuli et al. [13]	Classification and analysis of text transcription from Thai depression assessment tasks among patients with depression	The study explores the use of Encoder-Decoder transformer LLM such as XLM-RoBERTa for classification of depression by utilizing the text transcriptions of speech responses and Thai depression assessment questions	XLM-RoBERTa has a tendency to perform well in classification tasks, while only using small amount of data. Showing potential in helping with insufficient health care workers's problem and improve healthcare service accessibility.	The study concluded that speech-to-text transcription can be used for screening depression, but may not be effective for other psychiatric conditions or conditions that involve overlapping depressive symptoms.
Udomnuchaisup et al. [14]	Analysis of the 5Rs in Thailand Medication Error Classification through Natural Language Processing	The study explores the usage of pre-trained Large Language Models (PLMs) to enhance the detection and prevention of medication errors in hospitals, specifically expanding the coverage of the Five Rights (5R) error categorization system.	The supervised learning classification models that utilized sentence embeddings, still outperform zero-shot classification from PLMs due to insufficiency of pre-trained language models that leverage Thai language datasets from the general domain.	The study mentions the lack of PLMs employing Thai medical datasets and Thai context in general, within training. Which caused the model to underperform with Thai languages and even more so in medical-related context.
Noraset et al. [15]	Language-agnostic deep learning framework for automatic monitoring of population-level mental health from social networks	The researchers present LAPoMM as a framework for monitoring real-time mental health indicators from social media data without labeled datasets in low-resource languages.	LAPoMM leverages the capabilities of fine-tuned, language-independent, LaBSE model, to conduct mental signal classification and showed that extracted mental signals are in line with population-level mental health statistics	The study mentioned the lack of ability to deal with messages that are misinformed, inorganic, or automatically created, that caused the extracted social signals to be interpreted incorrectly.

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308731; this version posted June 11, 2024. The copyright holder has placed this preprint (which was not certified by peer review) in the Public Domain. It is no longer restricted by copyright. Anyone can legally share, reuse, remix, or adapt this material for any purpose without crediting the original authors.

Author	Study Title	Area/Focus	Outcomes	Challenges
Ponthongmak et al. [16]	Development and external validation of automated ICD-10 coding from discharge summaries using deep learning approaches	The aim of the study is to develop an automated international classification of diseases (ICD) coding tool using NLP and discharge summary texts from Thailand utilizing model trained on long English clinical text translated from Thai-English text data.	The CNN-PubMedBERT model performed second-best to PLM-ICD model, both performed well in the prediction of ICD-10 classification.	The study mentions that further external validation would be required before consideration as a clinically useful tool as Area Under the Precision-Recall Curve (AUPRCs) for these algorithms are still very modest and not sufficiently high enough for clinical use.
Suvirat et al. [17]	De-identification of Thai Free-text Clinical Notes	The research aims to explore, develop, and evaluate a method for de-identifying Thai free text clinical notes using language models.	The research utilized MEDPSU-RoBERTa, Multilingual BERT and Wangchan-BERTa, revealing that MEDPSU-RoBERTa performed best at highest score of all three models for de-identification task.	The authors present a promising results on the use of BERT architecture model for de-identification task. However, there can be other ethical requirements that needs to be fulfilled from utilizing LLMs to handle Personally Identifiable Information.
Boonrit et al. [18]	Evaluation of the Appropriateness of Thai-language Answers from ChatGPT for Questions About Medication Use in Common Illnesses	The study focused on validating the outputs from ChatGPT for answering questions related to medicines used for common sickness in Thai language.	The study shows that the responses were quite appropriate with some answers providing inappropriate information.	ChatGPT may respond in a way to make people think that they need to take certain medications, when in reality, they don't, this is considered a self-diagnosis and can be harmful.

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308731; this version posted June 11, 2024. The copyright holder has placed this preprint (which was not certified by peer review) in the Public Domain. It is no longer restricted by copyright. Anyone can legally share, reuse, remix, or adapt this material for any purpose without crediting the original authors.

6 DISCUSSION

This review highlights the potential benefits of utilizing Large Language Models (LLMs) in the healthcare sector in Thailand. Notably, existing studies have primarily utilized LLMs for classification tasks, with limited exploration of their potential for conversational interactions with patients. This bias may be attributed to the concerns surrounding LLMs, which need to be addressed.

Once the limitations of LLMs are mitigated, they can offer numerous benefits in the healthcare sector. For instance, LLMs can help alleviate the issues of slow traffic, long wait times, and accessibility barriers that often discourage patients from seeking medical attention. By enabling patients to initially diagnose their symptoms and engage in conversational interactions with chatbots, LLMs can provide personalized advice and guidance, potentially reducing the need for human verification.

Furthermore, LLMs can be integrated into healthcare facilities to enhance workflow efficiency, such as in inpatient diagnosis coding [12]. The emergence of multimodal LLMs, capable of processing text, visual, and auditory information, offers additional opportunities for diagnosis, such as visual symptom analysis and auditory examinations.

Overall, LLMs and Large Multimodal Models (LMMs) have the potential to revolutionize healthcare by providing accurate and natural language interactions.

However, this review also highlights several limitations and challenges that need to be addressed. The reliance of LLMs on training data can introduce biases and inaccuracies, particularly in lower-resource languages like Thai. Moreover, ethical concerns surrounding the use of LLMs in sensitive fields like healthcare necessitate careful examination and weighing of benefits and concerns before their widespread adoption. It is essential to examine and balance the benefits and concerns of utilizing LLMs in healthcare before they can be integrated into everyday practices.

6.1 Limitations of contextual understanding and Interpretability

As LLMs is simply large artificial neural networks, it learned to make decisions by utilizing a fixed weight and bias values that are tuned in the training process (as called parameters). The decision it used to generate text is simply based on the statistical patterns of the text, they lack the reasoning behind the decisions they made, it means that LLMs cannot understand the subtlety of medical diagnoses which can had multiple different factors to keep in mind for diagnosis, which means that LLMs lacks the true understanding of the underlying medical concepts and reasons, and so, LLMs may produce inaccurate information when met with lower-resources information such as rare diseases or complex cases. And because LLMs utilize patterns for generation, they lack actual reasoning behind the decisions that it chose, which make it difficult for healthcare professionals to trust and validate its recommendations, which is a problem in critical decision-making scenarios where it require a clear understanding of the reasoning behind the suggested diagnoses or treatment options.

6.2 Lack of resources and Reliance on training data

How well LLMs will perform heavily relies on the training data, since any biases and inaccuracies that present in the data will be learned by the model. And depending on the dataset, the model could perform worse when utilizing the LLMs with other languages such as Thai, due to context differences and low resources about the language in the training data. It is crucial to train the base model with knowledge specific to the field you are utilizing it, and include the languages you are utilizing it in, and not in the fine-tuning phase, as fine-tuning can degrade the performance of certain knowledge even more from the base model [7]. One can utilize RAG to create a ground-truth information for the LLMs to reference in their responses [11], allowing for frequent update to the information without retraining the entire model, and to prevent hallucinations.

6.3 Impact on healthcare professionals

Even though LLMs can be a useful tool to healthcare professionals in reducing works that need to be done by human, one cannot completely rely on the output of such models, as that can potentially result in the lack of critical thinking and independent clinical judgement. The usage of LLMs must be as a tool to aid in the works, not replacing the role entirely.

6.4 Ethical Considerations

One must address the concerns of using LLMs to handle sensitive information such as patient information, as privacy and security of the data must be prioritized to protect sensitive health information. Proper consent mechanisms should be established to ensure that the patients are aware of their data being processed by the language models. And any ethical concerns regarding the potential liability and accountability of artificial intelligence systems in healthcare need to be addressed to avoid potential harm to patients and caused legal implications.

6.5 Real-world testing

Before LLMs can be used in real-world applications, there is a need for rigorous validation and real-world testing to assess its performance and reliability of the outcomes. Further evaluations and clinical trials are necessary to understand the true effectiveness of LLMs in improving diagnosis accuracy, reducing errors and biases and enhancing patient care and workflow.

7 CONCLUSIONS

This review highlights the potential benefits of utilizing Large Language Models (LLMs) in the healthcare field in Thailand, including mitigating issues related to patient waiting times and accessibility, improving diagnosis accuracy, and enhancing workflow efficiency. However, several limitations and concerns must be addressed before widespread adoption, including the reliance on training data, which can introduce biases and inaccuracies, the lack of resources with Thai languages, and ethical concerns related to patient privacy and security. Furthermore, LLMs lack true understanding of medical concepts and reasoning, which can lead to inaccurate information and difficulties in trusting and validating their recommendations. To overcome

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308731; this version posted June 11, 2024. The copyright holder has placed this preprint (which was not certified by peer review) in the Public Domain. It is no longer restricted by copyright. Anyone can legally share, reuse, remix, or adapt this material for any purpose without crediting the original authors.

these challenges, rigorous validation and real-world testing are necessary to assess the performance and reliability of LLMs in healthcare applications. Additionally, addressing the limitations of contextual understanding, interpretability, and ethical considerations is crucial to ensure the effective and responsible integration of LLMs in healthcare. By acknowledging and addressing these challenges, LLMs can potentially become a valuable tool in improving healthcare outcomes and workflow efficiency in Thailand.

8 FUTURE CONSIDERATIONS

As the potential benefits of LLMs in healthcare become increasingly evident, it is essential to address the limitations and challenges associated with their adoption. To fully realize the benefits of LLMs, future research should focus on the following areas:

1. Addressing biases and inaccuracies: Developing strategies to mitigate the impact of biased training data on LLMs' performance, particularly in lower-resource languages like Thai, is crucial. This may involve developing more diverse and representative training datasets or implementing bias-reduction techniques, and collaborating with healthcare professionals and domain experts to help in tuning the datasets and implementing validation mechanisms.
2. Enhancing multimodal capabilities: Further research is needed to fully explore the potential of multimodal LLMs in healthcare, including the integration of visual, auditory, and other sensory information to improve diagnosis accuracy and patient outcomes.
3. Ethical considerations: The use of LLMs in healthcare raises important ethical concerns, including issues related to patient privacy, data security, and the potential for biased decision-making. Future research should prioritize the development of ethical guidelines and frameworks for the responsible deployment of LLMs in healthcare.
4. Human-LLM collaboration: To maximize the benefits of LLMs in healthcare, it is essential to develop effective strategies for human-LLM collaboration, ensuring that healthcare professionals and LLMs work together seamlessly to provide high-quality patient care.
5. Scalability and integration: Future research should focus on developing scalable and integrable LLM solutions that can be easily incorporated into existing healthcare systems, facilitating widespread adoption and minimizing disruption to healthcare services.

By addressing these key areas, we can unlock the full potential of LLMs in healthcare, improving patient outcomes, enhancing healthcare efficiency, and enhancing the healthcare sector as a whole.

REFERENCES

[1] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention Is All You Need, 2023.

[2] Ranwir K Sinha, Asitava Deb Roy, Nikhil Kumar, and Himel Mondal. Applicability of ChatGPT in assisting to solve higher order problems in pathology. *Cureus*, 15(2): e35237, 2 2023.

[3] Fatima Muftić, Merjem Kadunić, Almina Mušinbegović, and Ali Abd Almisreb. Exploring Medical Breakthroughs: A Systematic Review of ChatGPT Applications in Healthcare. *Southeast Europe Journal of Soft Computing*, 12(1):13–41, 5 2023. ISSN 2233–1859. doi: 10.21533/scjournal.v12i1.252. URL http://inquiry.ius.edu.ba/index.php/scjournal/article/view/252. Number: 1.

[4] G. E. Berrios and T. Dening. Pseudohallucinations: a conceptual history. *Psychological Medicine*, 26:753–763, 1996.

[5] Nick McKenna, Tianyi Li, Liang Cheng, Mohammad Javad Hosseini, Mark Johnson, and Mark Steedman. Sources of Hallucination by Large Language Models on Inference Tasks, 2023.

[6] Kunat Pipatanakul, Phatrasek Jirabovonvisut, Potsawee Manakul, Sittipong Sripaisarnmongkol, Ruangsak Patomwong, Pathomporn Chokchainant, and Kasima Tharnpipitchai. Typhoon: Thai Large Language Models, 2023.

[7] Sreyan Ghosh, Chandra Kiran Reddy Evuru, Sonal Kumar, Ramaneswaran S, Deepali Aneja, Zeyu Jin, Ramani Duraiswami, and Dinesh Manocha. A Closer Look at the Limitations of Instruction Tuning, 2024.

[8] Georgi Gerganov. llama.cpp. https://github.com/ggerganov/llama.cpp, 2023. URL https://github.com/ggerganov/llama.cpp.

[9] Chroma Core Team. Chroma. https://github.com/chroma-core/chroma, 2022. URL https://github.com/chroma-core/chroma.

[10] Etienne Dilocker, Bob van Luijt, Byron Voorbach, Mohd Shukri Hasan, Abdel Rodriguez, Dirk Alexander Kulawiak, Marcin Antas, and Parker Duckworth. Weaviate. https://github.com/weaviate/weaviate. URL https://github.com/weaviate/weaviate.

[11] Oded Ovadia, Menachem Brief, Moshik Mishaeli, and Oren Elisha. Fine-Tuning or Retrieval? Comparing Knowledge Injection in LLMs, 2024.

[12] Kerdkiat Suvirat, Detphop Tanasanchonnakul, Sawrawit Chairat, and Sitthichok Chaichulee. Leveraging Language Models for Inpatient Diagnosis Coding. *Applied Sciences*, 13(16), 2023. ISSN 2076-3417. doi: 10.3390/app13169450. URL https://www.mdpi.com/2076-3417/13/16/9450.

[13] Adirek Munthuli, Pakinee Pooprasert, Nittayapa Klangpornkun, Phongphan Phienphanich, Chutamanee Onsuwan, Kankamol Jaisin, Keerati Pattanaseri, Juthawadee Lortrakul, and Charturong Tantibundhit. Classification and analysis of text transcription from Thai depression assessment tasks among patients with depression. *PLOS ONE*, 18(3):e0283095, 3 2023. ISSN 1932-6203. doi: 10.1371/journal.pone.0283095. URL http://dx.doi.org/10.1371/journal.pone.0283095.

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308731; this version posted June 11, 2024. The copyright holder has placed this preprint (which was not certified by peer review) in the Public Domain. It is no longer restricted by copyright. Anyone can legally share, reuse, remix, or adapt this material for any purpose without crediting the original authors.

[14] Peachyasitt Udomnuchaisup, Aurawan Imsombut, Picha Suwannahitatorn, and Thammakorn Saethang. Analysis of the 5rs in Thailand Medication Error Classification through Natural Language Processing. In *2023 20th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, pages 281–284, 6 2023. doi: 10.1109/JCSSE58229.2023.10202096.

[15] Thanapon Noraset, Krittin Chatrinan, Tanisa Tawichsri, Tipajin Thaipisutikul, and Suppawong Tuarob. Language-agnostic deep learning framework for automatic monitoring of population-level mental health from social networks. *Journal of Biomedical Informatics*, 133:104145, 2022. ISSN 1532-0464. doi: https://doi.org/10.1016/j.jbi.2022.104145. URL https://www.sciencedirect.com/science/article/pii/S1532046422001587.

[16] Wanchana Ponthongmak, Ratchainant Thammasudjarit, Gareth McKay, John Attia, Nawanan Theera-Ampornpunt, and Ammarin Thakkinstian. Development and external validation of automated ICD-10 coding from discharge summaries using deep learning approaches. *Informatics in Medicine Unlocked*, 38:101227, 3 2023. doi: 10.1016/j.imu.2023.101227.

[17] Kerdkiat Suvirat, Sawrawit Chairat, Kanakorn Horsiritham, Chanon Kongkamol, and Sitthichok Chaichulee. De-identification of Thai Free-text Clinical Notes. In *2023 15th Biomedical Engineering International Conference (BMEiCON)*, pages 1–5, 10 2023. doi: 10.1109/BMEiCON60347.2023.10321968.

[18] Nantaphong Boonrit, Ashley M. Hopkins, and Warit Ruanglertboon. Evaluation of the Appropriateness of Thai-language Answers from ChatGPT for Questions About Medication Use in Common Illnesses. *Thai Journal of Pharmacy Practice*, 15(4):1134–1148, 9 2023. ISSN 3027-6136. URL https://he01.tci-thaijo.org/index.php/TJPP/article/view/264517. Original work published in Thai.