



HAL
open science

Automated Processing of Medication Error Reports with a GPT Transformer Model

Hanae Touati, Rafika Thabet, Franck Fontanili, Marc Pruski, Marie-Hélène Cleostrate, Marie-Noëlle Cufi, Elyes Lamine

► To cite this version:

Hanae Touati, Rafika Thabet, Franck Fontanili, Marc Pruski, Marie-Hélène Cleostrate, et al.. Automated Processing of Medication Error Reports with a GPT Transformer Model. 2024 IEEE/ACS - 21st International Conference on Computer Systems and Applications (AICCSA), Oct 2024, Sousse, Tunisia. pp.1-6, <10.1109/AICCSA63423.2024.10912561>. <hal-04998660>

HAL Id: hal-04998660

<https://imt-mines-albi.hal.science/hal-04998660v1>

Submitted on 10 Apr 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Automated Processing of Medication Error Reports with a GPT Transformer Model

Hanae Touati

Toulouse University, CGI, IMT Mines Albi
Albi, France
hanae.touati@mines-albi.fr

Rafika Thabet

Toulouse University, CGI, IMT Mines Albi
Sousse University, ISITCom,
MARS Laboratory, LR17RS05, Tunisia
Albi, France
rafika.thabet@mines-albi.fr

Franck Fontanili

Toulouse University, CGI,
IMT Mines Albi
Albi, France
franck.fontanili@mines-albi.fr

Marc Pruski

Atout Majeur Concept
Toulouse, France
mpruski@synora.fr

Marie-Hélène Cleostrate

Intercommunal Hospital Center
Castres-Mazamet, France
marie-helene.cleostrate@chic-cm.fr

Marie-Noëlle Cufi

Intercommunal Hospital Center
Castres-Mazamet, France
marie-noelle.cufi@chic-cm.fr

Elyes Lamine

Toulouse University, ISIS Engineering, INUC
Toulouse University, CGI, IMT Mines Albi
Castres, France
elyes.lamine@univ-jfc.fr

Abstract—Because of their potentially serious consequences, Medication errors (ME) represent a major challenge for healthcare facilities. To manage these errors and minimize their seriousness, healthcare professionals follow a collaborative management process that relies on reporting and then analyzing the reports filled in by them via various reporting tools, both digital and paper-based. These tools can be customized requiring specific information to be entered in multiple forms with commonly the presence of textual descriptions to fill in a free-text field. Therefore, text analysis is crucial for thoroughly understanding and effectively analyzing medication errors. Given the large volume of reports to be quickly processed, it is essential to help healthcare professionals prioritize which ME to analyze. In this context, we propose, in this work, processing ME reports with natural language processing tasks using Transformer models such as GPT. In this study, we present the extraction of key information from the reports to help structure textual descriptions of ME with the GPT-4 transformer model. The results obtained show the potential of this model to extract relevant information from ME descriptions in French language without any deep fine-tuning.

Index Terms—Risk Management, Natural Language Processing, Data extraction, Transformer Model, Medication Error, Medication Use Process.

I. INTRODUCTION

In the medical field and along the care pathway, medication errors (MEs) are among the critical risks affecting patient safety. They encompass any unintentional error made by a healthcare professional, patient, or any other person, occurring during the care process involving a medicine from prescription to administration and monitoring [1]. These errors can have varying levels of severity, compromising the effectiveness of care. Thus, managing these risks associated with the hospital

Medication Use Process (MUP) is important for improving patient safety and optimizing therapeutic outcomes.

To achieve this, in France, the management of those risks is organized across three levels [2]: (1) the internal level within Healthcare Facilities (HFs), (2) the regional level orchestrated by Regional Health Agencies (RHA), and Regional Support Structures (RSS), and (3) the national level represented by the High Authority for Health (HAS). Each of these levels plays a predominant role in (i) the analysis of medication error, (ii) the establishment of action plans, and (iii) the development of frameworks, standards, recommendations, and regulatory best practices aimed at minimizing the risks of MEs. However, despite these risks being managed at different levels, allowing for nationwide general management, there remains a significant number of risks that are not reported at either the regional or national levels and must be managed internally by the healthcare facility. Among these risks are medication errors (MEs), which can have serious consequences, and recur in different care units, but cannot all be managed at once. Therefore, in our work, we focus on the management of medication errors within the internal operations of healthcare facilities. Therefore, in our work, we focus on the management of ME within the internal level in french public or private HFs.

At this level, HFs have established a management process that includes reporting MEs, analyzing their causes and effects within the Experience Feedback Committee (EFC) meeting, and finally implementing and monitoring corrective and preventive action plans. However, the application of this process faces major obstacles, particularly the inadequacy of existing approaches for managing all MEs occurring across various

services and within the time constraints imposed by each facility’s internal regulations. This reality highlights the urgent need to find new approaches to support the EFC in their risk management sessions. Given this current manual and time-consuming approach, it is nearly impossible to effectively and quickly cover all reported MEs for a period occurring across different services during the same meeting. This situation raises a primary issue of how best to assist healthcare professionals in managing MEs internally at the hospital. Among the considered solutions, the automation of some steps of the ME management process, particularly when selecting MEs to analyze, emerges as a promising path. It offers the hope of a faster and more reliable response to this problem, ensuring optimal patient safety within healthcare facilities. To do so Artificial Intelligence (AI) and data science models have proved their transformative potential for enhancement within different fields. They can process vast amounts of data rapidly, allowing professionals to analyze several data in an optimized timing compared to the time required for the current methods. Furthermore, data science techniques, such as natural language processing (NLP), can be applied to identify patterns and trends in different forms.

From their proven capability to significantly transform data processing and analysis across various fields, we decide for the same objective of enhancing patient safety and care quality within HFs, to integrate AI and data science techniques into the analysis of MEs. By applying NLP and employing powerful transformer models such as GPT-4, we aim to delve deeper into the textual data of ME descriptions. Within this approach, we aim to rapidly extract important information and discern patterns that traditional methods might overlook. This initiative is driven by the desire to help healthcare professionals rapidly understand the ME occurring in the HF and establish more effective prevention strategies, ultimately leading to a more secure patient’s care pathway.

In this work, we present the use of NLP tasks to extract several important information from ME reports. In Section II, we present the general context of our work, namely the internal hospital ME management process, which relies on analyzing descriptive textual reports. The textual form of those reports presents an initial challenge to automating this process. To address this, we will present AI techniques that have emerged in the medical field and can help structure ME reports. Finally, the section concludes with the presentation of language Transformer models. Section III presents the application of the GPT-4 model on ME reports for understanding and structuring MEs. Lastly, we present our conclusions and perspectives in Section IV.

II. ME MANAGEMENT BACKGROUND AND AI APPLICATIONS

In this section, we explore the current methodology for managing risks related to the MUP at the internal level of the HFs. Then we present the the application of AI within the medical field, highlighting their relevance and impact. Finally,

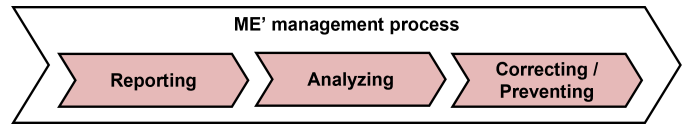


Fig. 1. Internal process of ME management in hospitals

we discuss the transformative potential of Transformer models architecture, examining the main existing models.

A. Internal management process of medication errors in hospitals

In France, the management of MEs internally in hospital systems follows a process primarily composed of three interdependent steps as shown in Figure 1. Each step plays an essential role in addressing efficiently and effectively those errors. In the following, we introduce the three steps:

- **Reporting:** In this initial step, healthcare professionals report each medication error by filling out a report form, which can be either paper-based or digital, depending on the facility’s preferences. The form can include predefined fields used to gather essential information describing the ME or a non-structured free textual field, defined according to the preferences of each HF.
- **Analyzing:** In this step, an EFC composed of different healthcare professionals reviews previously reported errors to understand their causes and effects. Based on this analysis, they develop corrective and preventive action plans aiming to prevent future occurrences of similar errors. They convene two to three times a year for meetings that last no more than two hours each.
- **Correcting/preventing:** In this stage, the corrective and preventive action plans developed during the analysis phase are put into practice. The implementation of these actions is continuously monitored to ensure they are effective.

This structured process helps address individual errors. It also contributes to a culture of continuous improvement and safety within hospital systems, ultimately enhancing patient care quality and safety. However, it suffers from several limitations, mainly related to the dependency on the reporting phase.

In this context, the management revolves around reports in reporting systems designed to collect specific information characterizing each ME. These systems often require the reporter (healthcare professional) to fill out forms detailing the error, combining textual descriptions with structured elements such as checkboxes, drop-down lists, or pre-filled fields set differently depending on the HFs and their specialties. These descriptions help explain the context and circumstances of the error, while the structured elements facilitate understanding, classification, and statistical analysis of MEs. Nonetheless, this process can suffer from under-reporting due to staff workload, fear of sanction, or lack of feedback following previous reports.

Moreover, the effectiveness of analyzing the reports can be limited by the quality of the information provided, which varies depending on the reporter’s ability to effectively communicate the details of the error, lack of time, or other factors. After a detailed study of existing practices and a comparison of various management strategies of different hospitals in France, we have identified that the major obstacle lies in the heterogeneity of the reporting tools used within health establishments, which vary considerably from one establishment to another. This diversity, resulting from the customization of digital or paper reporting systems according to the preferences of each hospital, leads to a lack of standardization in the information collected. Furthermore, the tendency to favor free text entry and the absence of mandatory fields in these tools result in non-structured and often incomplete reports, thus limiting the accuracy, depth of error analysis, and the potential for using data science techniques for automated analysis.

The unstructured free text in reports makes it difficult to extract consistent and relevant data for analysis. Consequently, the EFC, responsible for analyzing the root causes of ME and implementing a corrective and preventive action plan, often has to refer to supplementary documents, such as patient records or service and report sheets, to get an overview of the circumstances surrounding each error. This does not promote a comprehensive understanding of MEs, thus hindering efforts and time to improve patient safety and quality of care. To overcome this obstacle, the solution is to structure the reporting phase using a standard form, to harmonize the collection of essential data for both the reporting and analysis of MEs. But before such a form is validated and used, there is an enlarge volume of ME reports to be utilized to improve the management process. The commonality among all these reports is the descriptive text. Thus, this improvement will primarily rely on processing and extracting relevant information from this descriptive text.

B. AI application in the medical field for language processing

The integration of AI in healthcare represents a significant advancement, promising to enhance the quality and efficiency of medical care [3], in terms of diagnostics, prevention, and outcomes. The application of AI in processing health data has shown substantial performance, especially in the analysis of medical imaging, disease prediction, and treatment personalization [4]. However, it is in NLP that AI can demonstrate revolutionary potential for the medical sector [5], particularly in managing different risks related to the MUP, such as Adverse events or adverse drug events [6], [7]. Although MEs are a major concern for patient safety, involving significant risks to their health, effective management of these errors requires meticulous analysis of the collected data, namely the ME reports which are often in the form of complex descriptive texts. So, we want to test the potential of AI, specifically NLP, to achieve this goal.

Powered by deep learning models and advanced algorithms, NLP systems are capable of understanding context, extracting relevant information, and processing natural language in a way that reflects the complexity of human communications in all

languages [8]. These systems enable text classification, sentiment analysis, named entity recognition, text generation, and more, across all languages and in various contexts, even those as specific as the medical field [9]. The ability to interpret specific medical language allows for more effective analysis of medical data and offers solutions for better improvement. Additionally, in the case of MEs, the application of AI in natural language processing paves the way for more intuitive and accessible ME reporting and tracking systems, as shown in many other researches such as [10]and [11].

The promises of NLP in the medical field encompass improvements in clinical documentation, optimization of communication between healthcare professionals and patients, and support for medical research [12]. By enabling rapid and precise analysis of vast volumes of text, NLP helps unlock the potential of unstructured health data, opening new perspectives for evidence-based care and personalized medicine.

C. Transformer Models for NLP Tasks

“Transformer” models have revolutionized the NLP with their unique architecture that includes attention mechanisms [13]. They are capable of learning from a large volume of unlabelled data in various languages and using this knowledge to perform different NLP tasks by considering the context of the text studied. When applied to ME descriptions, NLP tasks can analyze complex and technical language specifications in the medical field to extract key information characterizing MEs from unstructured descriptive texts [14]. The level of knowledge they have acquired during their training allows them to detect correlations and trends that would elude manual analyses or traditional statistical methods. Thus, Transformer models can help structure the textual data with Named entity recognition (NER) as presented by [15] and more detailed by [16] or with Retrieval-augmented generation (RAG) [17]. The advent of Transformer models marked a revolution in the field of automated NLP, establishing new foundations for the understanding and generation of language by machines [9]. These models are distinguished by their ability to process data sequences in parallel and to capture complex dependencies between words, thanks to their attention mechanism, which allows weighting the relative importance of different words in a sentence [18]. Since the introduction of BERT [19], Transformer models have rapidly evolved, with each new model bringing significant improvements in terms of text understanding and generation. Table I presents a comparison between the most well-known and used Transformer models since their first appearance. The comparison in the table includes the architecture of the model, whether it follows an encoder, decoder, or encoder-decoder structure, and the NLP tasks it can perform (text and token classification, document summarization or translation, question answering, text generation, or completion of incomplete texts). Today, there are several pre-trained Transformer models designed to execute specific NLP tasks in specific languages. Generally, they are ready to be used on data in English but can be adapted and trained for the desired language. For example, the study by

[20] refined BERT for named entity recognition from clinical data in German, [21] used AraBERT [22], a new model derived from BERT, for a named entity recognition task in Arabic dialect texts.

In the context of enhancing the extraction of key information from ME reports in French, a challenge arises due to the limited availability of sufficiently large datasets to fine-tune a transformer model specifically for this purpose. To overcome this obstacle, we have opted to leverage the pre-existing knowledge of a model that understands French. GPT-4, the latest version of GPT model known for its robust performance in language comprehension and generation across multiple languages, was chosen. By adapting GPT-4, already pre-trained in a multitude of tasks and languages including French, we customize its application to specifically address the nuances of extracting critical data from ME reports. This strategic approach can not only save significant resources and time that would otherwise be spent on developing a new model but also could ensure a high level of accuracy and efficiency in processing ME reports in French.

III. METHODOLOGY FOR PROCESSING MEDICATION ERROR REPORTS

For the extraction of pertinent information from ME reports using the GPT-4 model, we followed the many steps as presented in 2. Our approach integrates unsupervised and supervised extraction to efficiently process and analyze the complex data derived from various healthcare reporting tools. In this section, we detail each phase of the process, from data collection to the application of advanced named entity recognition algorithms, illustrating how these methods assess our innovative data analysis strategy.

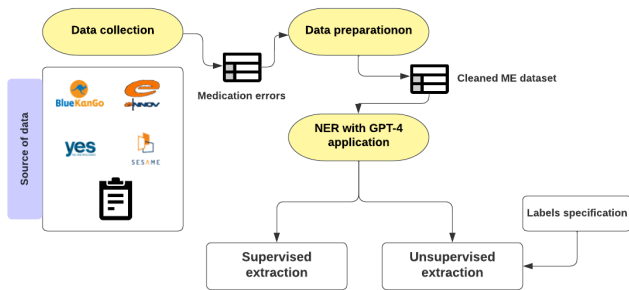


Fig. 2. Methodology for processing Medication Errors descriptions

A. Data Collection

The data collection for this study was based on ME reports from various HFs. These reports were gathered using several tools used in France, including paper forms and digital platforms such as BlueKango, Sesame, Yes and Ennov. Each tool provides a reporting guide that prompts users to respond to multiple-choice questions, targeting specific characteristics of ME. Despite the variety of formats, all these tools share a

common feature: they allow for the inclusion of a free text description of the incident. This uniformity in the collection of textual data paves the way for consistent analysis, despite the diversity of sources and reporting formats.

B. Extraction of information

Given the challenge of processing and analyzing vast quantities of free descriptive texts from ME reports, adopting an AI-based approach was crucial. Specifically, employing a Transformer model like GPT-4 was identified as the most effective method for extracting relevant information from these complex descriptions without prior annotation. In our study, we leveraged the capabilities of GPT-4, with its ability to understand and generate natural text, to analyze ME report examples and extract key information through Named Entity Recognition (NER). This approach allows for the automatic identification and classification of key information within the texts, such as the types of drugs involved, the stages of the care process where the error occurred, and the observed consequences, according to the data cited in a report. This exploration was conducted in two distinct Phases:

- **Phase 1: Unsupervised Extraction**

In the first phase, we utilized the model to perform named entity recognition without pre-specifying which labels to look for. The model demonstrated a remarkable ability to identify various types of entities within the texts, such as dates, medication names, consequences, as well as the identities of healthcare professionals and/or patients involved. It's important to note that the nature and quantity of information extracted varied significantly from one report to another, reflecting the diversity and inconsistency of the data provided in each report.

- **Phase 2: Supervised Extraction**

In the second phase, we refined our approach by explicitly specifying the labels we were searching for medication, date, time, healthcare professional, patient, cause, and consequence. This specificity served as a guide for the model during the extraction of information corresponding to these categories, thus enhancing the relevance and depth of the data collected.

C. Results and discussion

Figure 3 presents the results obtained within the supervised and unsupervised significant information related to ME extractions with the GPT-4 model. It compares unsupervised and supervised information extraction using GPT-4 from an ME description. The unsupervised method identifies basic details like dates, medication, and involved healthcare professionals. In contrast, with precise labels, the supervised approach extracts more detailed and structured information including exact dates, times, medication details, involved people, and the consequences of the error. This highlights the advantage of supervised learning in providing more comprehensive and contextually relevant information from descriptive text.

The model demonstrated its capability to discern a variety of relevant labels from the descriptive texts. While the accuracy is

TABLE I
COMPARATIVE TABLE OF TRANSFORMER MODELS

Model	Architecture			Tasks					
	Encoder	Decoder	Encoder & Decoder	Classification	Summarising	Translation	Q/R	Text generation	Fill Mask
BERT [19]	✓			Text & Token			✓		✓
GPT [23]		✓		Text & Token	✓	✓	✓	✓	✓
ALBERT [24]	✓			Text & Token			✓		✓
RoBERTa [25]	✓			Text & Token			✓		✓
T5 [26]			✓	Text & Token	✓	✓	✓	✓	✓
BART [27]			✓	Text & Token	✓	✓	~✓		~✓
XLNet [28]		✓		Text & Token			✓		✓

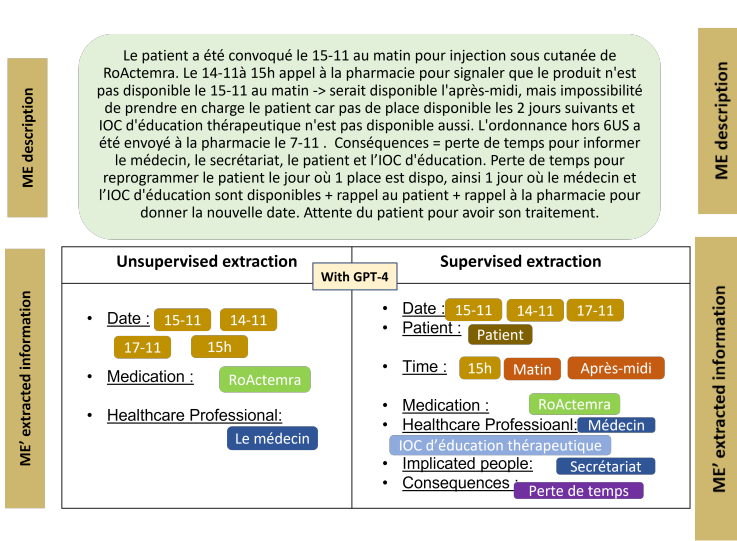


Fig. 3. Relevant ME information extraction with GPT-4

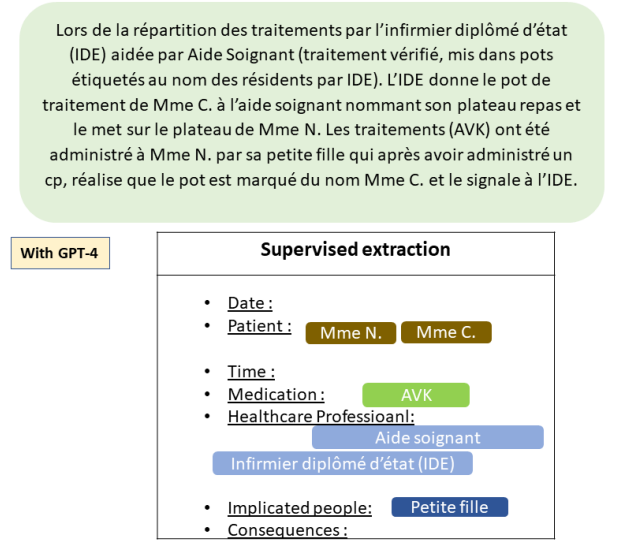


Fig. 4. Supervised ME information extraction with GPT-4

not perfect at 100% and the extractions require manual verification or completion, the model establishes a solid foundation for the automatic labeling of reports. These initial successes indicate that the use of Transformer models for text analysis in the case of ME reports is a promising solution.

It is important to note that the information to extract depends on what is mentioned in the reports, as demonstrated by the examples. In the first test case in Figure 3, the model successfully identified the date, medication, time, involved healthcare professionals, the patient, the consequences, and even other involved people if they exist. However, in the second test case, presented in Figure 4, the report was less detailed. So the model was able to recognize only the entities concerning the involved people (patient, healthcare professional, and others), and the medication. Which illustrates the dependency of the extraction's effectiveness on the information initially present in the reports. Despite including causes as entities to be extracted, the model did not extract any because they were not specified in the reports. This underscores a limitation inherent to the approach which strictly depends on the input data provided.

Finally, our results demonstrate our approach's robustness in harnessing the GPT-4 model for structuring textual data

and extracting relevant information, underscoring its significant applicability in medical data in the French language. This method enhances the interpretability and accessibility of complex medical texts, enabling more accurate and actionable insights in the field.

IV. CONCLUSION

In conclusion, this article highlighted the complexity and challenges associated with managing MEs in healthcare facilities, emphasizing the importance of effectively and rapidly processing reports provided by healthcare professionals. Through examining the current process, which largely relies on the analysis of free textual descriptions collected via various tools, we have identified an urgent need for innovation to overcome the challenges posed by the large volume of reports and time constraints.

The automatic processing of these reports, through natural language processing (NLP) and specifically the use of Transformer models like GPT, appears as a promising solution. This approach offers the ability to quickly extract key information necessary for the prioritization and management of MEs, and it also facilitates an efficient structuring of data, allowing

for more in-depth analysis and a better understanding of the reported incidents. Moving forward, in our future work we focus on automatic validation and leveraging the extracted information to label medication error reports, thus facilitating the transition to the classification of MEs. This will assist healthcare professionals in the deeper analysis and categorization of these errors, enhancing the prediction and correction of the occurred errors. This next phase will involve refining our NLP models to handle the variability and complexity of ME descriptions more effectively, ensuring robustness and reliability in real-world applications. This strategic use of AI and machine learning will improve the accuracy and speed of response in healthcare facilities and also support a proactive stance in patient safety and quality care initiatives.

ACKNOWLEDGMENT

The authors would like to express their heartfelt appreciation to the Occitanie Region of France and the Atout Majeur Concept (AMC) company for funding and supporting this project. Special thanks are due to Mr. Dimitri Duval, the CEO of AMC for their unwavering support and invaluable contributions of expertise and resources throughout this research.

REFERENCES

- [1] E. Schmitt, D. Antier, C. Bernheim, E. Dufay, M. C. Husson, and E. Tissot, "Dictionnaire français de l'erreur médicamenteuse," *Société française de pharmacie clinique*, 2006.
- [2] OMS, "Haute Autorité de Santé: Déclarer les événements indésirables graves associés aux soins (EIGS)," 2022.
- [3] G. Damiani, G. Altamura, M. Zedda, M. C. Nurchis, G. Aulino, A. H. Alizadeh, F. Cazzato, G. Della Morte, M. Caputo, S. Grassi *et al.*, "Potentiality of algorithms and artificial intelligence adoption to improve medication management in primary care: a systematic review," *BMJ open*, vol. 13, no. 3, p. e065301, 2023.
- [4] S. V. G. Subrahmanya, D. K. Shetty, V. Patil, B. Z. Hameed, R. Paul, K. Smriti, N. Naik, and B. K. Somani, "The role of data science in healthcare advancements: applications, benefits, and future prospects," *Irish Journal of Medical Science (1971-)*, vol. 191, no. 4, pp. 1473–1483, 2022.
- [5] S. A. Alowais, S. S. Alghamdi, N. Alsuhebany, T. Alqahtani, A. I. Alshaya, S. N. Almohareb, A. Aldairem, M. Alrashed, K. Bin Saleh, H. A. Badreldin *et al.*, "Revolutionizing healthcare: the role of artificial intelligence in clinical practice," *BMC medical education*, vol. 23, no. 1, p. 689, 2023.
- [6] A. Fong, N. Harriott, D. M. Walters, H. Foley, R. Morrissey, and R. R. Ratwani, "Integrating natural language processing expertise with patient safety event review committees to improve the analysis of medication events," *International journal of medical informatics*, vol. 104, pp. 120–125, 2017.
- [7] C. Liang, S. Zhou, B. Yao, D. Hood, and Y. Gong, "Toward systems-centered analysis of patient safety events: Improving root cause analysis by optimized incident classification and information presentation," *International Journal of Medical Informatics*, vol. 135, p. 104054, 2020.
- [8] K. Chowdhary and K. Chowdhary, "Natural language processing," *Fundamentals of artificial intelligence*, pp. 603–649, 2020.
- [9] A. Gillioz, J. Casas, E. Mugellini, and O. Abou Khaled, "Overview of the transformer-based models for nlp tasks," in *2020 15th Conference on computer science and information systems (FedCSIS)*. IEEE, 2020, pp. 179–183.
- [10] A. B. Chapman, K. S. Peterson, P. R. Alba, S. L. DuVall, and O. V. Patterson, "Detecting adverse drug events with rapidly trained classification models," *Drug safety*, vol. 42, pp. 147–156, 2019.
- [11] L. Chen, Y. Gu, X. Ji, Z. Sun, H. Li, Y. Gao, and Y. Huang, "Extracting medications and associated adverse drug events using a natural language processing system combining knowledge base and deep learning," *Journal of the American Medical Informatics Association*, vol. 27, no. 1, pp. 56–64, 2020.
- [12] S. Locke, A. Bashall, S. Al-Adely, J. Moore, A. Wilson, and G. B. Kitchen, "Natural language processing in medicine: a review," *Trends in Anaesthesia and Critical Care*, vol. 38, pp. 4–9, 2021.
- [13] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.
- [14] J. Rebanc, I. Samsten, and P. Papapetrou, "Exploiting complex medical data with interpretable deep learning for adverse drug event prediction," *Artificial Intelligence in Medicine*, vol. 109, p. 101942, 2020.
- [15] C. Berragan, A. Singleton, A. Calafiore, and J. Morley, "Transformer based named entity recognition for place name extraction from unstructured text," *International Journal of Geographical Information Science*, vol. 37, no. 4, pp. 747–766, 2023.
- [16] A. Goyal, V. Gupta, and M. Kumar, "Recent named entity recognition and classification techniques: a systematic review," *Computer Science Review*, vol. 29, pp. 21–43, 2018.
- [17] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W.-t. Yih, T. Rocktäschel *et al.*, "Retrieval-augmented generation for knowledge-intensive nlp tasks," *Advances in Neural Information Processing Systems*, vol. 33, pp. 9459–9474, 2020.
- [18] A. Rahali and M. A. Akhloufi, "End-to-end transformer-based models in textual-based nlp," *AI*, vol. 4, no. 1, pp. 54–110, 2023.
- [19] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, 2018.
- [20] J. Frei and F. Kramer, "Annotated dataset creation through large language models for non-english medical nlp," *Journal of Biomedical Informatics*, vol. 145, p. 104478, 2023.
- [21] M. Jarrar, M. Khalilia, and S. Ghanem, "Wojood: Nested arabic named entity corpus and recognition using bert," *arXiv preprint arXiv:2205.09651*, 2022.
- [22] W. Antoun, F. Baly, and H. Hajj, "Arabert: Transformer-based model for arabic language understanding," *arXiv preprint arXiv:2003.00104*, 2020.
- [23] A. Radford, K. Narasimhan, T. Salimans, I. Sutskever *et al.*, "Improving language understanding by generative pre-training," 2018.
- [24] Z. Lan, M. Chen, S. Goodman, K. Gimpel, P. Sharma, and R. Soricut, "Albert: A lite bert for self-supervised learning of language representations," *arXiv preprint arXiv:1909.11942*, 2019.
- [25] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov, "Roberta: A robustly optimized bert pretraining approach," *arXiv preprint arXiv:1907.11692*, 2019.
- [26] C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, and P. J. Liu, "Exploring the limits of transfer learning with a unified text-to-text transformer," *Journal of machine learning research*, vol. 21, no. 140, pp. 1–67, 2020.
- [27] M. Lewis, Y. Liu, N. Goyal, M. Ghazvininejad, A. Mohamed, O. Levy, V. Stoyanov, and L. Zettlemoyer, "Bart: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension," *arXiv preprint arXiv:1910.13461*, 2019.
- [28] Z. Yang, Z. Dai, Y. Yang, J. Carbonell, R. R. Salakhutdinov, and Q. V. Le, "Xlnet: Generalized autoregressive pretraining for language understanding," *Advances in neural information processing systems*, vol. 32, 2019.