Accelerating Pharmacovigilance using Large Language Models

M V Sai Prakash Ganesh Parab Meghana Veeramalla
Siddartha Reddy Saisubramaniam Gopalakrishnan Varun V
Vishal Pagidipally Vishal Vaddina

Applied Research, Quantiphi

The 17th ACM International Conference on Web Search and Data Mining

Abstract

Pharmacovigilance involves collecting, analyzing, and reporting data on adverse events, allowing for informed regulatory decisions. Manual systems face challenges in handling data volume, potentially leading to oversight and delays. Automation with advanced technologies can be a practical solution to mitigate these challenges and ensure efficient data management.

The potential application of Large Language Models (LLMs) in pharmacovigilance is explained, starting with an overview of the process covering all stages of the lifecycle. The pivotal process of scrutinizing documents for relevant adverse effects is emphasized, along with measures to enhance effectiveness. The strategy for generating informative summaries, specifically focusing on adverse effects and their antecedent occurrences, is delineated. The imperative requirement for factual accuracy validation through the implementation of a fact-checking mechanism is highlighted. The framework's efficacy is demonstrated with a focus on output fidelity and summary informativeness. Quantifiable evidence of the benefits of the method is provided, advocating for the adoption of this framework in pharmacovigilance. The conclusion addresses potential refinements.

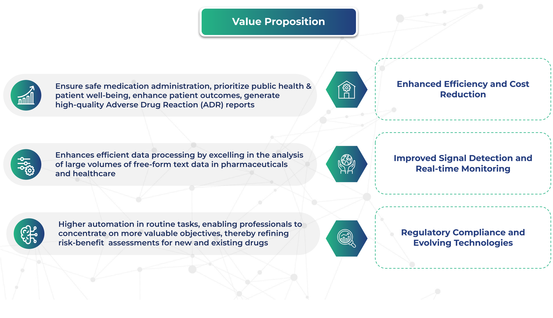
Pharmacovigilance

Pharmacovigilance (PV) is the science and activities related to the detection, assessment, understanding, and prevention of adverse effects or any other drug-related problems [1].

- Ensures patient safety by monitoring and evaluating drug safety throughout its life cycle.
- Detects and minimizes the risks associated with pharmaceutical products.
- Contributes to the overall improvement of public health.

Motivation

The adoption of AI is transforming traditional manual processes, resulting in reduced time and resource consumption. One significant trend is the increasing use of Natural Language Processing (NLP) and machine learning for automated data extraction and analysis [2]. In terms of signal detection and real-time monitoring, AI's proficiency in analyzing large datasets motivates its application for quicker identification of safety concerns, with a growing trend of integrating AI-driven monitoring systems for real-time surveillance.



Capabilities of Our Solution

Document Filtering
- Preprocessing the documents and generating chunks
- Filter out documents without adverse effects
- Store those with adverse effects in ElasticSearch

Subject-based Summarization
- Summarizing content based on specific subjects or topics
- Extract key information and generate summaries related to a particular subject, irrespective of the doc source

Document ID-based Summarization
- Summarizing content based on Doc-IDs (User Input : Doc ID)
- The system retrieves the document associated with the doc ID
- Generates a summary based on the content found in the specified document

Question Answering
- Question Answering System that operates atop the provided documents
- Semantic Retrieval of documents through query-based mechanisms
- Context referencing in the generated responses to enhance confidence levels

Architecture

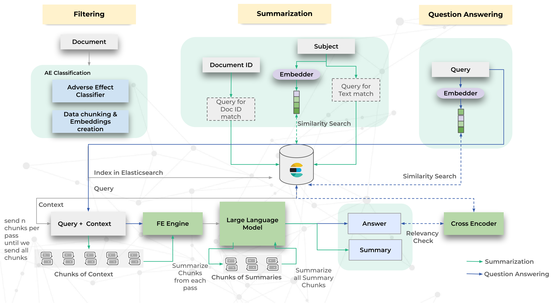


Figure 1. Comprehensive Architecture of the Solution

Document Filtering

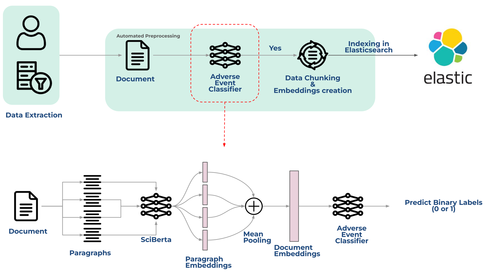


Figure 2. Documents undergo filtration via a Neural Network Classifier to remove non-adverse event documents, while those with adverse effects are stored

Question Answering and Summarization

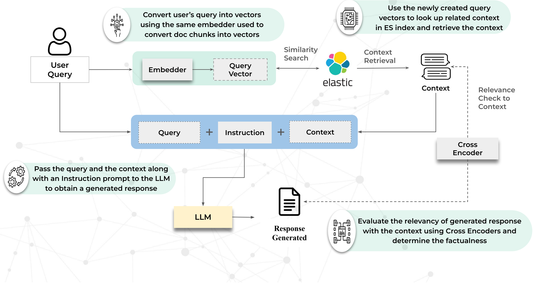


Figure 3. Architectural design of a Question Answering framework integrated with adverse event documents, leveraging Large Language Models

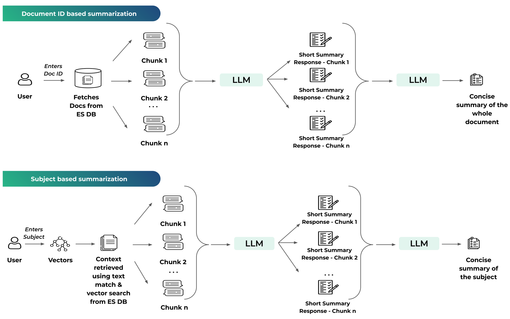


Figure 4. The architecture comprises Subject-based Summarization and Document-based Summarization modules, enabling users to generate summaries either for individual documents or for specific subjects (drugs), tailored to their requirements

Validation and Results

Reader Evaluation

Question
What are the potential risks and side effects associated with the use of azithromycin?

Response
The potential risks and side effects associated with the use of azithromycin may include serious allergic reactions(^a), liver damage(^a, ^d), serious heart rhythm changes(^a), cardiovascular issues like arrhythmias(^b), gastrointestinal issues like constipation(^b), vomiting(^b), diarrhea(^c), and nausea(^c). Other possible side effects could include skin reactions(^b), chest pain(^c), dizziness(^c), rash(^c), changes in heartbeat(^d), and muscle weakness(^d). Based on the given facts, it can be inferred that azithromycin tablets may have a range of potential risks and side effects.

Rationale
a. PID-1 states that azithromycin tablets can cause serious side effects, including serious allergic reactions, liver damage, and serious heart rhythm changes.

b. PID-2 states that adverse reactions reported with azithromycin may include allergic reactions, cardiovascular issues like arrhythmias, gastrointestinal issues like constipation and vomiting, and skin reactions.

c. PID-3 states that the most common treatment-related side effects of azithromycin in adult patients include gastrointestinal issues like diarrhea and nausea, as well as other side effects like chest pain, dizziness, and rash.

d. PID-5 states that potential risks and side effects of azithromycin include liver problems, changes in heartbeat, muscle weakness, and breathing problems.

e. No other facts provide information about the potential risks and side effects of azithromycin.

f. Based on the information from the given facts, it can be inferred that azithromycin tablets may have a range of potential risks and side effects, including serious allergic reactions, liver damage, heart rhythm changes, gastrointestinal issues, skin reactions, changes in heartbeat, and muscle weakness.

Faithfulness Score: 0.97652

Retriever Evaluation

	@ 1	@ 3	@ 5
MRR	0.843	0.812	0.776
NDCG	0.934	0.917	0.874

References

[1] P. Beninger.
Pharmacovigilance: an overview.
Clinical therapeutics, 40(12):1991–2004, 2018.

[2] A. J. Thirunavukarasu, D. S. J. Ting, K. Elangovan, L. Gutierrez, T. F. Tan, and D. S. W. Ting.
Large language models in medicine.
Nature medicine, 29(8):1930–1940, 2023.