Harnessing Transformers for Detecting Adverse Drug Reaction and Customized Causality Explanation using Generative AI

Ankan Bera   
*Data & Analytics, Advanced Analytics*   
*PwC India*ankan.bera@pwc.com

Raktim Chakraborty  
Data & Analytics, Advanced Analytics  
*PwC India*   
rakitm.chakraborty@pwc.comRik Das  
Data & Analytics, Advanced Analytics  
*PwC India*   
rik.das@pwc.com

Indranil Mitra  
Data & Analytics, Advanced Analytics  
*PwC India*   
indranil.mitra@pwc.comSayantani Ghosh  
Data & Analytics, Advanced Analytics  
*PwC India*  
sayantani.g.ghosh@pwc.com

Prasun Nandy  
Data & Analytics, Advanced Analytics  
*PwC India*  
prasun.nandy@pwc.com

*Abstract*— Pharmacovigilance plays an important role in monitoring safety of pharmaceutical products. There is an abundance of unstructured data available in social media and online reviews that can provide valuable insights into adverse drug reaction (ADR). This research paper aims to explore the efficient techniques of pharmacovigilance in identifying ADRs from unstructured data available in social media and reviews. This research work proposes the use of a transformer-based model for ADR classification, followed by named entity recognition (NER) to identify the causality between drugs and adverse effects. Furthermore, the output of NER is provided as a customized prompt to the generative AI model to generate a user interpretable explanation. Enhancing the accuracy of NER is carried out using the SciBERT model, which is specifically designed to capture drug and effect names. By utilizing this combination of techniques, the research aims to improve the efficiency and effectiveness of pharmacovigilance processes, enabling earlier detection and reporting of ADRs. To our knowledge, previous works have not explored the application of generative AI based customized explanation of causality. Patients, healthcare professionals, regulatory agencies, and pharmaceutical companies can benefit from the timely and accurate identification of ADRs, allowing for improved drug safety monitoring, better decision-making regarding drug usage, and the development of proactive measures to mitigate risks.

Keywords— Pharmacovigilance, Transformers, BERT, GPT, Generative AI, Text Classification, Named Entity Recognition, ADR

# Introduction

Pharmacovigilance is a critical component of the drug development process that focuses on the monitoring, evaluation, and prevention of adverse drug reactions (ADRs) caused by pharmaceutical products. Pharmacovigilance clients include pharmaceutical companies, regulatory authorities, healthcare professionals, patients, and consumers. It is an essential practice for ensuring the safety and efficacy of drugs in the market [1]. However, traditional methods of pharmacovigilance rely on manual reporting, which can lead to underreporting and delays in identifying ADRs, hence there is an immediate need for complete or partial automation of the process [2].

The current state of pharmacovigilance reflects a growing acknowledgment of its significance within healthcare systems globally. Numerous countries have developed strong pharmacovigilance frameworks that prioritize the methodical collection and analysis of adverse event data. Initiatives are underway to enhance reporting systems, encourage active participation from healthcare professionals and patients in reporting adverse events, and facilitate improved data sharing among regulatory agencies. Furthermore, the adoption of advanced technologies and data analytics is gaining momentum, leading to more efficient detection of signals and assessment of risks. This can help the stakeholders take proactive measures, such as issuing warnings or recalls, improving patient safety, and minimizing legal risks [3].

The use of deep learning techniques in automated pharmacovigilance has the potential to improve the accuracy and efficiency of ADR detection to a large extent [4]. Recent advancements in transformer-based models, such as BERT and GPT, have shown promising results in understanding and processing natural text inputs. Fine-tuning the pre-trained transformer models with drugs and adverse drug reaction related information can significantly improve the performance of ADR detection and to identify the concerned drugs and their adverse effects for causality identification.

The challenge, therefore, is to design a system which will accurately and quickly identify ADRs from unstructured data and present an automated alternative to complement the current process.

The objectives of this study are as follows:

• To evaluate the effectiveness of transformer-based models in ADR classification from natural text inputs.

• To identify causality from the classified ADR data with the help of named entity recognition

• To create customized prompts from NER and classification output for generating causality explanation with the help of generative AI.

The results of this study can have significant impact for the pharmaceutical industry by providing a more efficient and accurate approach to ADR detection, leading to improved patient safety, public trust, and better understanding of drug effectiveness.

The structure of this research paper includes a literature survey, research methodology, results and analysis, and conclusion. The literature survey reviews previous research on pharmacovigilance and deep learning techniques for ADR detection. The research methodology describes the data collection process, deep learning models used, their evaluation metrics. The results and analysis section present the experimental results and provides an analysis of the findings which includes a comparative study of the performance of different transformer models and identify causality. Finally, the conclusion summarizes the research findings, highlights the contributions, and suggests future research directions.

# Literature Survey

Deep learning techniques have emerged as promising tools for ADR identification due to their ability to learn complex patterns from large-scale data. In this literature review, we explore the use of deep learning methods for ADR identification and highlight key findings from relevant studies.

Wang et al. [5] proposed a deep neural network model for detecting potential ADRs. They utilized electronic health records (EHRs) and presented efficacy of their model in identifying ADRs with high accuracy and precision. Their study showcased the potential of deep learning in leveraging structured medical data for ADR detection. Dey et al. [6] presented an interpretable deep learning framework for predicting ADRs. Their model incorporated attention mechanisms to identify important features and provided explanations for its predictions. This interpretable approach offers valuable insights into the factors contributing to ADRs, aiding clinicians and researchers in understanding the underlying mechanisms. Ensemble deep learning methods were employed by Christopoulou et al. [7] for ADR and medication relation extraction from EHRs. Their study demonstrated that combining multiple deep learning models improved the performance of ADR identification. The ensemble approach enhances the robustness and generalization of ADR detection systems. Cocos et al. [8] focused on utilizing recurrent neural network architectures to label ADRs in Twitter posts. Their deep learning model effectively extracted information from social media data and showed potential for real-time pharmacovigilance. Social media mining provides a valuable source of information for ADR monitoring due to its wide reach and quick dissemination of user experiences. Li et al. [9] proposed an end-to-end deep learning model for extracting information related to ADRs from EHR notes. Their model leveraged the power of deep learning in natural language processing and demonstrated competitive performance compared to traditional methods. Deep learning models enable automated analysis of unstructured clinical text, facilitating efficient ADR detection. Rezaei et al. [10] investigated the detection of ADRs in social media using deep learning methods. Their study demonstrated the potential of deep learning for ADR monitoring in online platforms. Mining social media data provides a valuable complement to traditional pharmacovigilance systems, enabling the detection of ADRs that might not be captured through other channels. Machine learning-based identification and rule-based normalization of ADRs in drug labels were explored by Tiftikci et al. [11]. They developed a machine learning model combined with rule-based approaches to accurately identify and normalize ADRs. Their study emphasized the importance of integrating different techniques for comprehensive ADR analysis. A comprehensive survey by Nguyen et al. [12] provided an overview of ADR studies, including data sources, tasks, and machine learning methods. The authors highlighted the potential of deep learning in ADR identification, particularly in the analysis of EHRs and social media data. The survey intends to be a valuable resource for researchers interested in ADR detection using machine learning. Nikfarjam et al. [13] employed sequence labeling with word embedding cluster features for mining ADR mentions from social media. Their study demonstrated the effectiveness of deep learning in extracting ADR-related information from user-generated content. By harnessing the power of word embeddings, deep learning models can capture semantic relationships and improve ADR detection accuracy. Alimova and Tutubalina [14] developed an automated system for ADR detection from social media posts using machine learning techniques. Their study highlighted the potential of machine learning, including deep learning, However, a research gap identified in these studies is the limited use of transformer-based models in ADR classification. In addition, there is a research gap in identifying causality for ADRs using named entity recognition. Furthermore, no literature is identified showcasing the use of generative AI for customized explanation of causality. The authors have identified the aforementioned limitations in state-of-the-art and have addressed all of them using transformer based adverse drug reaction classification followed by causality identification using named entity recognition and generating causality explanation using generative AI.

# Research Methodology

ADR corpus classification is envisioned using transformers since the transformer architecture has replaced the traditional recurrent and convolutional neural network components of NLP models with a self-attention mechanism that allows the model to focus on different parts of the input sequence during each layer of computation.

In text classification, the transformer model takes as input a sequence of word embeddings, which are learned representations of words in a high-dimensional space. The model then applies a series of self-attention and feedforward layers to the input sequence, producing a final output vector that is passed through a SoftMax layer to obtain the predicted class probabilities. The self-attention mechanism in the Transformer model can be expressed mathematically as follows,

(1)

where Q, K, and V are matrices representing the queries, keys, and values, respectively, and dk is the dimensionality of the keys.

The self-attention mechanism calculates a weighted sum of the values, where the weights are determined by the dot product of the queries and keys after normalization. This allows the model to focus on different parts of the input sequence based on their relevance to the current context.

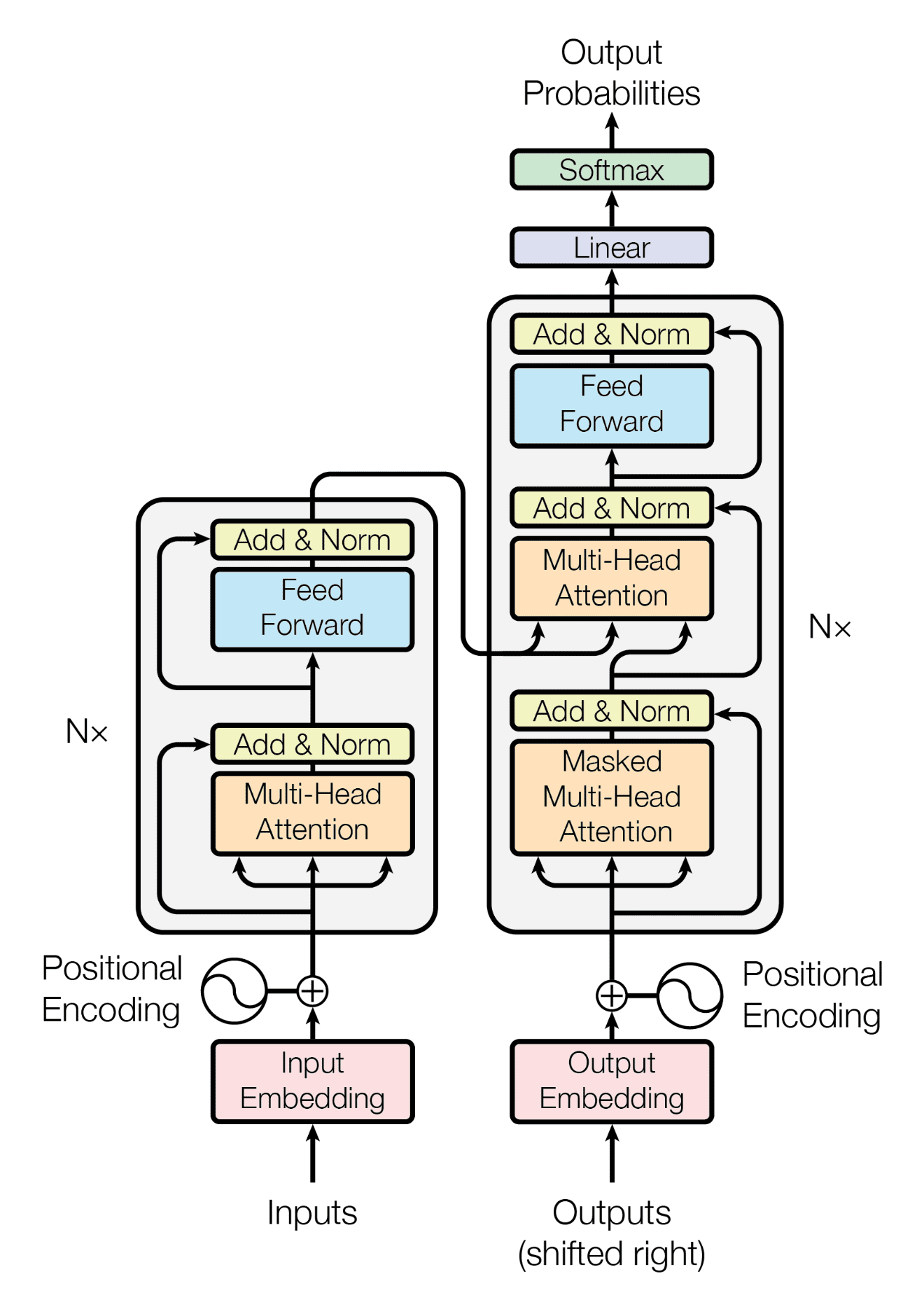


Fig 1: The encoder-decoder process of the Transformer architecture

For this experiment, we have used three pre-trained transformer models DistilBERT, xlm-RoBERTa and GPT-NEO-125M for text classification. DistilBERT and xlm-RoBERTa are two state of the art transformer models developed over the BERT model. DistilBERT is a smaller and efficient version of BERT while xlm-RoBERTa is a larger and more powerful BERT model that has shown promising performance on a wide range of NLP tasks. GPT-NEO-125M is a GPT based transformer model trained with 125 million parameters. While GPT models are primarily designed for language modelling and text generation, it has also shown impressive performance regarding text classification tasks.

Additionally, we have used a fine-tuned version of SciBERT model for performing named entity recognition with regards to identifying drug names and effects from natural text inputs. SciBERT is pre-trained on a large corpus of scientific publications, including abstracts and full-text articles from various domains such as computer science, biology, chemistry, and physics. The NER and classification output is used to design custom prompts for generating causality explanation using generative AI.

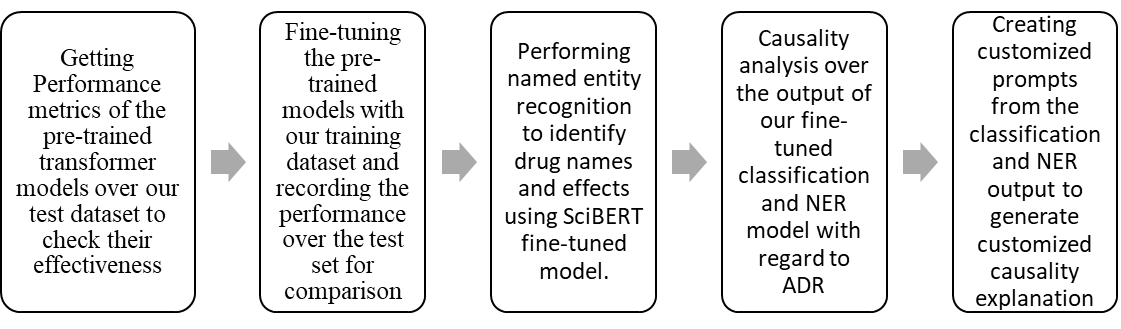


Fig 2: Process flow diagram of the experiment process

## Dataset Used: Ade\_corpus\_v2

In this study, we have used the Ade-corpus-v2-classifcation dataset, comprising of 23.5K labelled records classified as ADE or not, along with the relation extraction between ADE and drug [15]. The records in the form of natural texts, were collected from various online forums and medical journals regarding mentions of drugs and their effects on patients. Among the 23.5K records, 17.5K records have been used for fine-tuning our transformer models and rest of the 6K records were used for evaluating the performance metrics. The ADE-Corpus-V2 dataset is a collection of data on Adverse Drug Reactions (ADR). The dataset also provides information about the relationships between drugs and adverse effects through two files, DRUG-AE.rel and DRUG-DOSE.rel. These files provide data on the relationships between drugs and adverse effects and dosages respectively. Additionally, the dataset contains a file named ADE-NEG.txt that includes all the sentences in the ADE corpus that do not have any drug-related adverse effects.

In the case of ADR classification, high precision is desirable to avoid false positives, which can result in serious consequences for patient safety. However, high recall is also important to avoid false negatives, which can result in missed opportunities for early detection and prevention of ADRs. This work has used evaluation metrics including accuracy(A), precision(P), recall(R), and F1 score for which the equations are given below,

(2)

(3)

(4)

(5)

Where, TP = true positive,

TN = true negative,

FP = false positive,

FN = false negative.

# Results and Analysis

From our experiments, we have tried to assess the performance of all three fine-tuned versions of the pre-trained models with regards to the specific task of classifying a natural text as ADR related. We have also analysed the fine-tuned models’ performance based on their training and validation loss data. Finally, we have explored identification of causality relationship based on the output of the classification and the NER task which has been done using the fine-tuned SciBERT model.

## Evaluation the fine-tuned models in terms of detecting ADR

1. Performance metrics of the Fine-Tuned Models in Detecting ADR

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model Name** | **Accuracy (A)** | **Precision (P)** | **Recall (R)** | **F1-Score** |
| DistilBERT-finetuned | 0.95151 | 0.95275 | 0.95151 | 0.95188 |
| xlm-RoBERTa-finetuned | 0.95048 | 0.95212 | 0.95048 | 0.95094 |
| GPT-Neo-125M-finetuned | 0.93223 | 0.93625 | 0.93223 | 0.93188 |

Table I shows the performance metrics of the three fine-tuned models. From the result, we can observe that both the BERT models after fine-tuning has slightly outperformed the GPT-NEO-125M model in all the four metrics. This result can be attributed to the bi-directional architecture of BERT which can process the entire sequence of tokens in both the directions, allowing it to capture contextual information from the entire token sequence, while GPT’s left-to-right (unidirectional) architecture is not for classification tasks compared to BERT. Additionally, BERT is designed to be fine-tuned on specific downstream tasks, which results in more adaptability in learning for particular tasks such as ADR detection.

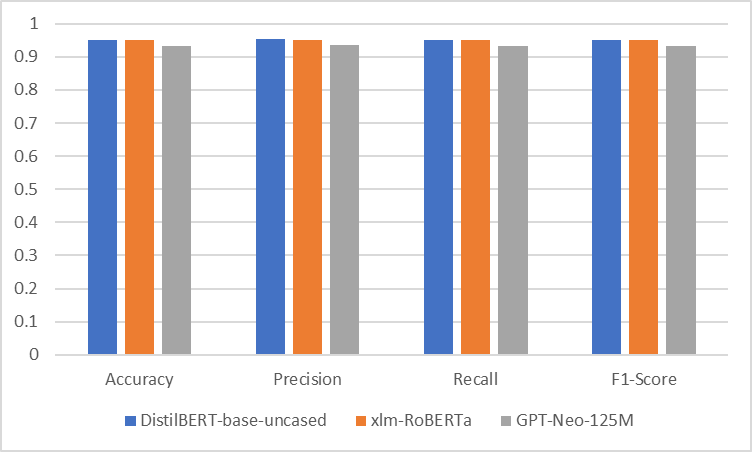


Fig 3: Graphical Representation of Table I

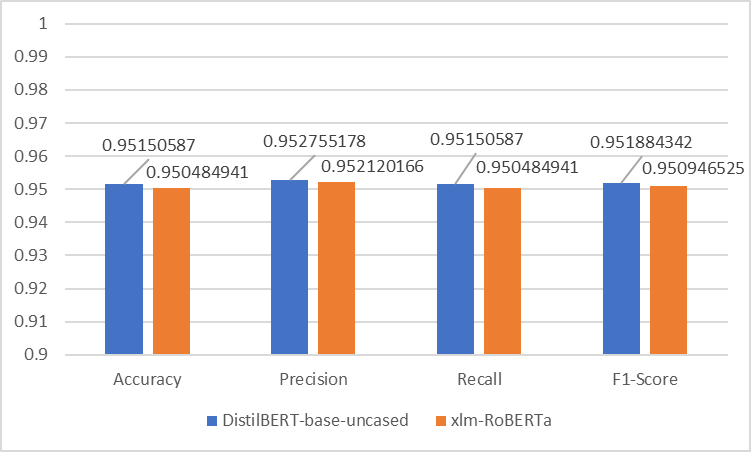


Fig 4: Graphical Comparison Between DistilBERT and xlm-RoBERTa after fine-tuning for ADR classification

Moreover, we can also observe from Figure 5, that both DistilBERT and xlm-RoBERTa have performed similarly in terms of accuracy and precision for ADR (Adverse Drug Reaction) classification. However, there is a slight difference in their recall and F1 score. DistilBERT has achieved an accuracy of 0.951, precision of 0.952, recall of 0.951, and F1 score of 0.952. In contrast, xlm-RoBERTa has achieved an accuracy of 0.950, precision of 0.952, recall of 0.950, and F1 score of 0.951. DistilBERT has performed slightly better in terms of recall and F1 score as compared to xlm-RoBERTa. Recall is a metric that measures the ability of the model to identify all the relevant instances of the positive class, whereas F1 score combines precision and recall to offer a holistic evaluation of the model's performance. With regards to ADR detection, we consider Recall to be a more important metric as the objective is not to miss even a single ADR event to ensure patient safety even if that comes at the cost of increased number of false positives.

The slightly better performance of DistilBERT is attributed to its architecture, which is a smaller and more efficient version of the original BERT model. DistilBERT employs a technique called distillation, in which the model is trained to mimic the behavior of the larger BERT model with fewer parameters. As a result, DistilBERT is faster and more memory-efficient while still achieving comparable performance to the original BERT model. In contrast, xlm-RoBERTa is based on the larger and more computationally expensive RoBERTa model. While xlm-RoBERTa has the capability to capture intricate patterns in the data owing to its increased capacity, it might necessitate a larger training dataset and lengthier training durations to attain optimal performance.

## Evaluating Model Performance with regards to Training and Validation Loss

Training and validation loss are two metrics used to analyze the performance of a transformer model during training. Training loss measures the error of the model's predictions on the training data during each epoch of training. It is determined by assessing the disparity between the model's predicted output and the actual output through the utilization of a loss function. The goal of training is to minimize the training loss, which means the model is becoming better at fitting the training data. Validation loss, on the other hand, measures the error of the model's predictions on a separate validation dataset that is not used for training. The validation loss is calculated in the same way as the training loss, but it provides an estimate of how well the model will generalize to new, unseen data. The goal of validation is to assess the model’s performance during training and detect signs of overfitting, where the model becomes excessively tailored to the training data and does not generalize well to new data.

Fig 6 shows the training loss curves plotted against the number of training epochs. As can be seen, the training loss steadily decreases for each epoch for xlm-RoBERTa and GPT-NEO-125M model, while for DistilBERT it spikes a little during epoch 5 but overall decreases for each epoch. GPT-NEO-125M can be seen to have 0 training loss from 7th epoch onwards, pointing to overfitting of the model to the training data, which in turn can be attributed to the comparatively poorer performance with the test data.

1. Training Loss per Epoch

|  |  |  |  |
| --- | --- | --- | --- |
| **Epoch** | **DistilBERT-finetuned** | **xlm-RoBERTa-finetuned** | **GPT-NEO-125M-finetuned** |
| **1** | 0.2657 | 0.1443 | 0.2437 |
| **2** | 0.1311 | 0.1344 | 0.1091 |
| **3** | 0.08 | 0.098 | 0.0454 |
| **4** | 0.0578 | 0.0705 | 0.0155 |
| **5** | 0.0377 | 0.0644 | 0.0027 |
| **6** | 0.0255 | 0.0401 | 0.0008 |
| **7** | 0.0194 | 0.0306 | 0.0001 |
| **8** | 0.0164 | 0.0182 | 0 |
| **9** | 0.0123 | 0.0139 | 0 |
| **10** | 0.0101 | 0.006 | 0 |



Fig 5: Graphical Represantation of Table II

1. Validation Loss per Epoch

|  |  |  |  |
| --- | --- | --- | --- |
| **Epoch** | **DistilBERT-finetuned** | **xlm-RoBERTa-finetuned** | **GPT-NEO-125M-finetuned** |
| **1** | 0.182549 | 0.278437 | 0.16493 |
| **2** | 0.169476 | 0.204062 | 0.183392 |
| **3** | 0.191256 | 0.190753 | 0.268707 |
| **4** | 0.271159 | 0.256249 | 0.355276 |
| **5** | 0.297953 | 0.312553 | 0.372711 |
| **6** | 0.301318 | 0.316171 | 0.383945 |
| **7** | 0.368504 | 0.370061 | 0.393788 |
| **8** | 0.376636 | 0.356444 | 0.400225 |
| **9** | 0.35497 | 0.403485 | 0.404048 |
| **10** | 0.361976 | 0.400605 | 0.405174 |

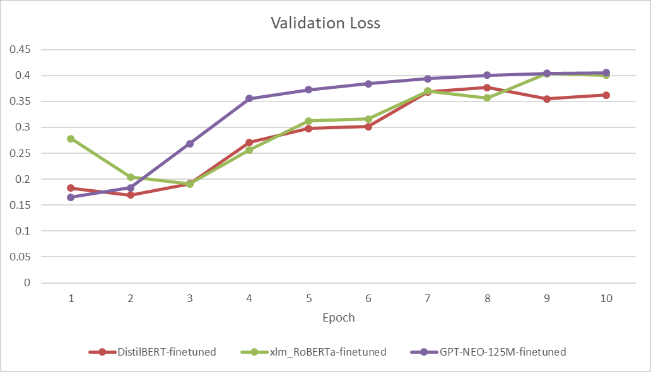


Fig 6: Graphical Representation of Table III

Fig 7 shows the validation loss curves plotted against the number of training epochs. It can be seen that the validation loss to be steadily increasing for GPT-NEO-125M model, in-contrast to the two BERT models where the graph shows validation loss to be decreasing for some epochs. Which indicates that the models are successfully learning from the training data during these epochs as opposed to memorizing the training data pattern and leading to overfitting, as can be seen for GPT-NEO-125M. Both the training loss and validation loss information points towards overfitting of the GPT-NEO-125M model and provides a valid reason for the poorer performance metrics compared to DistilBERT and xlm-RoBERTa.

1. Accuracy per Epoch

|  |  |  |  |
| --- | --- | --- | --- |
| **Epoch** | **DistilBERT-finetuned** | **xlm-RoBERTa-finetuned** | **GPT-NEO-125M-finetuned** |
| **1** | 0.931281 | 0.939105 | 0.932131 |
| **2** | 0.942507 | 0.937915 | 0.942847 |
| **3** | 0.950672 | 0.94659 | 0.94761 |
| **4** | 0.943018 | 0.941146 | 0.942167 |
| **5** | 0.949481 | 0.950162 | 0.946079 |
| **6** | 0.948971 | 0.944889 | 0.946079 |
| **7** | 0.942337 | 0.944378 | 0.94676 |
| **8** | 0.94744 | 0.951522 | 0.94659 |
| **9** | 0.949311 | 0.94676 | 0.94676 |
| **10** | 0.948801 | 0.949481 | 0.94676 |

Table IV displays the accuracy of the three fine-tuned models for each epoch. It can be seen that GPT-NEO-125M to be having a lower accuracy compared to the other two, supporting the conclusion of overfitting from the training and validation loss data.

## Causality Identification by Classification and NER result

After classifying a text as ADR related, for further analysis we have identified the concerned drugs and the adverse effect that have been mentioned in the text input using NER. Finally, the output of these two tasks have been used to create a custom prompt for generating customized causality explanation using generative AI. Fig 7 and 8 shows the output for an ADR and a non-ADR scenario respectively.

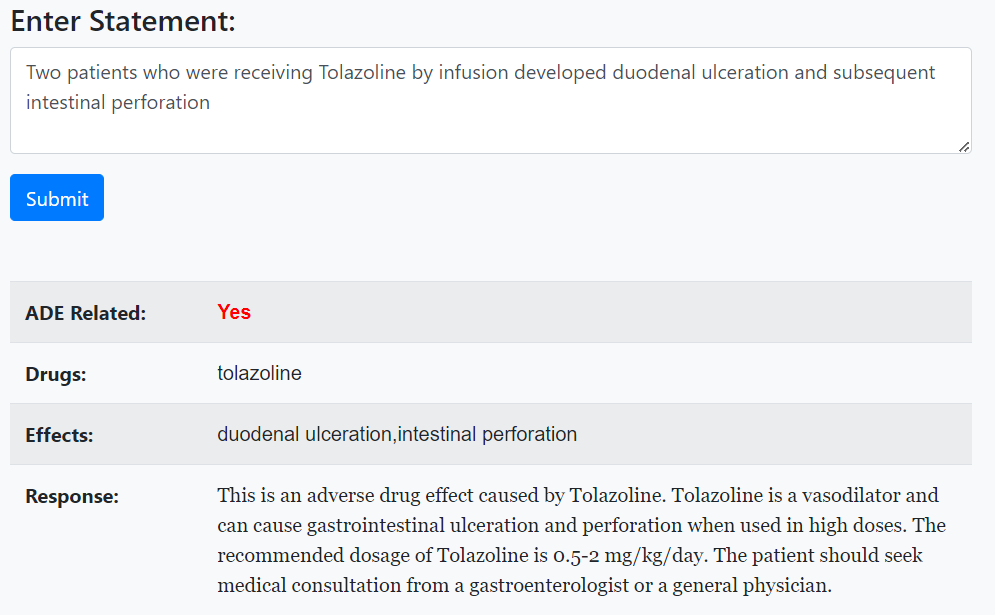


Fig 7: ADR-Related statement with explanation

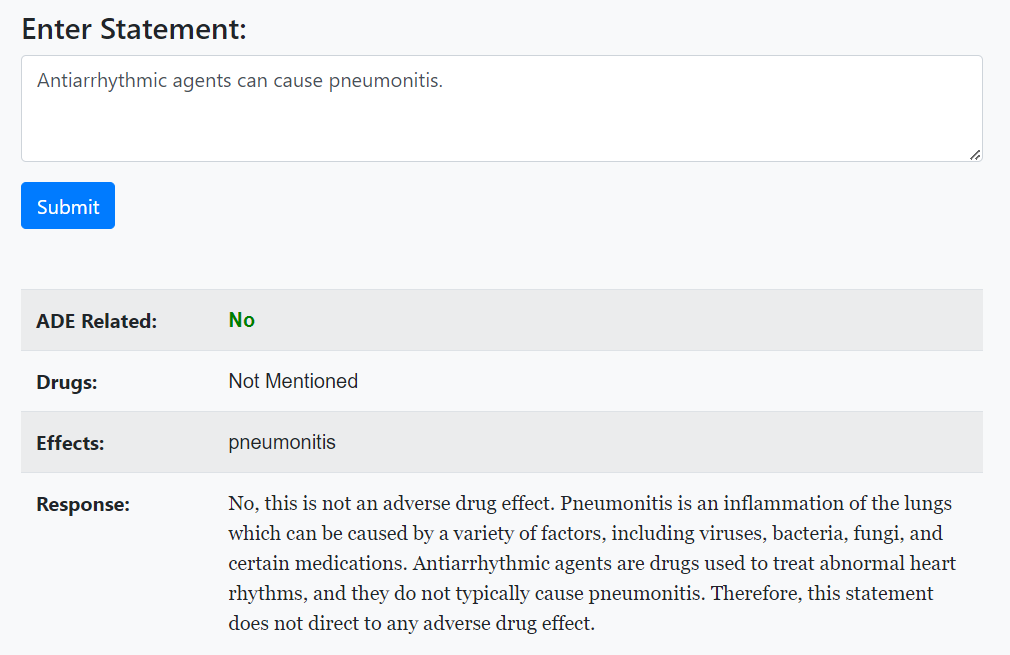


Fig 8: non-ADR-Related statement with explanation

# Conclusion

Pharmacovigilance plays a vital role in identifying and monitoring adverse drug reactions (ADRs), ensuring the safety and efficacy of pharmaceutical products. This research has demonstrated the significance of pharmacovigilance in utilizing advanced deep learning techniques to analyze unstructured data from social media and reviews, enabling the identification of ADRs that may otherwise go unnoticed. The research employed the Ade\_corpus\_v2 dataset, a pre-annotated dataset, for training three transformer models: DistilBERT, xlm-RoBERTa, and GPT-Neo-125M. Among these models, the finetuned DistilBERT model exhibited the best performance. Alternatives to pharmacovigilance include post marketing surveillance, clinical trials and spontaneous reporting system to detect suspected adverse drug events. Pharmacovigilance, as a specialized field, offers several advantages over its alternatives in terms of ensuring drug safety and monitoring adverse events. Firstly, compared to post-marketing surveillance alone, pharmacovigilance provides a proactive approach by systematically collecting and analysing data on drug safety throughout the product lifecycle. This enables the early detection of potential risks and the implementation of appropriate risk management strategies. Secondly, compared to clinical trials, which have limitations in terms of sample size and duration, pharmacovigilance captures real-world data from a broader patient population, including diverse demographics and co-morbidities. This allows for a more comprehensive understanding of drug safety profiles. Additionally, compared to spontaneous reporting systems alone, pharmacovigilance combines various data sources and analytical techniques, including signal detection algorithms and data mining, to identify safety signals more effectively. As the volume of unstructured data from social media and online forums continue to grow, there is a need for ongoing development and refinement of deep learning models to improve ADR identification and monitoring with pharmacovigilance. Additionally, further progress in generative AI for customized response generation will improve explainability of ADR identified using pharmacovigilance leading to informed decision making in healthcare sector.

##### VI. References

[1] N. Moore, D. Berdaï, P. Blin, and C. Droz, “Pharmacovigilance – The next chapter,” *Therapies*, vol. 74, no. 6, pp. 557–567, Dec. 2019, doi: 10.1016/j.therap.2019.09.004.

[2] J. Almenoff *et al.*, “Perspectives on the Use of Data Mining in Pharmacovigilance,” *Drug Saf.*, vol. 28, no. 11, pp. 981–1007, 2005, doi: 10.2165/00002018-200528110-00002.

[3] F. Ali *et al.*, “An intelligent healthcare monitoring framework using wearable sensors and social networking data,” *Futur. Gener. Comput. Syst.*, vol. 114, pp. 23–43, Jan. 2021, doi: 10.1016/j.future.2020.07.047.

[4] M. Salas *et al.*, “The Use of Artificial Intelligence in Pharmacovigilance: A Systematic Review of the Literature,” *Pharmaceut. Med.*, vol. 36, no. 5, pp. 295–306, Oct. 2022, doi: 10.1007/s40290-022-00441-z.

[5] C.-S. Wang, P.-J. Lin, C.-L. Cheng, S.-H. Tai, Y.-H. Kao Yang, and J.-H. Chiang, “Detecting Potential Adverse Drug Reactions Using a Deep Neural Network Model,” *J. Med. Internet Res.*, vol. 21, no. 2, p. e11016, Feb. 2019, doi: 10.2196/11016.

[6] S. Dey, H. Luo, A. Fokoue, J. Hu, and P. Zhang, “Predicting adverse drug reactions through interpretable deep learning framework,” *BMC Bioinformatics*, vol. 19, no. S21, p. 476, Dec. 2018, doi: 10.1186/s12859-018-2544-0.

[7] F. Christopoulou, T. T. Tran, S. K. Sahu, M. Miwa, and S. Ananiadou, “Adverse drug events and medication relation extraction in electronic health records with ensemble deep learning methods,” *J. Am. Med. Informatics Assoc.*, vol. 27, no. 1, pp. 39–46, Jan. 2020, doi: 10.1093/jamia/ocz101.

[8] A. Cocos, A. G. Fiks, and A. J. Masino, “Deep learning for pharmacovigilance: recurrent neural network architectures for labeling adverse drug reactions in Twitter posts,” *J. Am. Med. Informatics Assoc.*, vol. 24, no. 4, pp. 813–821, Jul. 2017, doi: 10.1093/jamia/ocw180.

[9] F. Li, W. Liu, and H. Yu, “Extraction of Information Related to Adverse Drug Events from Electronic Health Record Notes: Design of an End-to-End Model Based on Deep Learning,” *JMIR Med. Informatics*, vol. 6, no. 4, p. e12159, Nov. 2018, doi: 10.2196/12159.

[10] Z. Rezaei, H. Ebrahimpour-Komleh, B. Eslami, R. Chavoshinejad, and M. Totonchi, “Adverse Drug Reaction Detection in Social Media by Deepm Learning Methods.,” *Cell J.*, vol. 22, no. 3, pp. 319–324, Oct. 2020, doi: 10.22074/cellj.2020.6615.

[11] M. Tiftikci, A. Özgür, Y. He, and J. Hur, “Machine learning-based identification and rule-based normalization of adverse drug reactions in drug labels,” *BMC Bioinformatics*, vol. 20, no. S21, p. 707, Dec. 2019, doi: 10.1186/s12859-019-3195-5.

[12] D. A. Nguyen, C. H. Nguyen, and H. Mamitsuka, “A survey on adverse drug reaction studies: data, tasks and machine learning methods,” *Brief. Bioinform.*, vol. 22, no. 1, pp. 164–177, Jan. 2021, doi: 10.1093/bib/bbz140.

[13] A. Nikfarjam, A. Sarker, K. O’Connor, R. Ginn, and G. Gonzalez, “Pharmacovigilance from social media: mining adverse drug reaction mentions using sequence labeling with word embedding cluster features.,” *J. Am. Med. Inform. Assoc.*, vol. 22, no. 3, pp. 671–81, May 2015, doi: 10.1093/jamia/ocu041.

[14] I. Alimova and E. Tutubalina, “Automated Detection of Adverse Drug Reactions from Social Media Posts with Machine Learning,” 2018, pp. 3–15. doi: 10.1007/978-3-319-73013-4\_1.

[15] “ade\_corpus\_v2 · Datasets at Hugging Face.” https://huggingface.co/datasets/ade\_corpus\_v2/viewer/Ade\_corpus\_v2\_classification/train (accessed May 25, 2023).