AUTOMATED HEALTHCARE NLP FOR DETECTING ADVERSE DRUG REACTIONS

By Team 21

INTRODUCTION

Adverse Drug Reactions (ADRs) are harmful and unintended effects caused by medications, and identifying them early is essential for improving patient safety and drug monitoring.

However, most ADR information exists in unstructured text such as clinical notes, medical records, and research papers, making manual analysis slow and inefficient.

With advancements in Natural Language Processing (NLP), it is now possible to automatically analyze and classify biomedical text.

PROBLEM STATEMENT

This project aims to build an automated NLP-based system to detect Adverse drug reaction-related sentences using machine learning and modern transformer-based models, and to compare their performance in accurately identifying ADR information from medical text.

BASELINE APPROACH

Collect and prepare dataset : Use the ADE Corpus V2 dataset and perform necessary preprocessing such as cleaning , tokenizing, and converting labels into machine-readable form.

Implement a Baseline Model : Built a traditional ML classifier using TF-IDF vectorization and Logistic Regression to establish benchmark performance.

Apply Transformer-based Model : Fine-tune a pretrained DistilBERT model specifically for ADR sentence classification to leverage contextual language understanding.

Perform Hyperparameter Tuning : Experiment with different training settings, primarily learning rates, to identify the most effective configuration.

Evaluate and comapre Models : Compare baseline and transformer results using accuracy, precision, recall, and F1 score to determine performance improvement.

Conduct Error Analysis : Review misclassified examples to understand model limitations and identify areas for improvement, such as handling negation or implied ADR relationships.

LITERATURE REVIEW

Early approaches for detecting Adverse Drug Reactions relied on keyword matching and rule-based systems, but these methods struggled with context and variations in medical language.

Later, machine learning models such as **Logistic Regression and TF-IDF** improved automation but still lacked deep understanding of sentence meaning.

With the introduction of transformer-based models like BERT and DistilBERT, performance significantly improved because these models understand context rather than just isolated words. This project compares a traditional TF-IDF + Logistic Regression baseline with a fine-tuned DistilBERT model for ADR detection.

DATASET DESCRIPTION

The dataset used in this project is ADE Corpus V2, a publicly available benchmark dataset for ADR detection. The dataset consists of biomedical sentences sourced from drug safety literature, each labeled as either:

- ADR — the sentence explicitly or implicitly contains an adverse reaction
- Non-ADR — sentence does not describe harm caused by medication

DATASET STATISTICS:

ATTRIBUTE	VALUE
TOTAL SENTENCES	~4,272
ADR SENTENCES	~1,200
NON-ADR SENTENCES	~3,000

Steps used in data preprocessing:

- Text normalization
- Tokenization using the pretrained DistilBERT tokenizer
- Label encoding
- Train/Validation/Test split (80/10/10)

METHODOLOGY

DistilBERT is a lightweight compressed version of **BERT** retaining **~97%** of its performance while being faster and **computationally efficient. The model was fine-tuned for binary classification using HuggingFace Transformers.**

TRAINING PARAMETERS:

PARAMETER	VALUE
EPOCHS	3
BATCH SIZE	TRAIN : 16 EVAL : 32
ADR SENTENCES	~1,200
NON-ADR SENTENCES	~3,000

PROPOSED TIMELINE

PHASE	KEY GOALS AND DELIVERABLES	START DATE	END DATE
Planning	identify dataset availability and select modelling approach	September 29,2025	October 07, 2025
NLP development	Process dataset,build ML model and train transformer based model	October 07,2025	October 25,2025
Analysis	Compare model performance, perform error analysis and evaluate using metrics	Novermber 03,2025	November 15,2025
Delivery	Compile reseach report and finalize presentation	November 16,2025	November 24,2025

RESULTS

QUANTITATIVE EVALUATION

MODEL	ACCURACY	PRECISION	RECALL	F1-SCORE
TF-IDF +LR	0.88	0.88	0.76	0.84
DistilBert	0.9507	0.905	0.916	0.91

HYPERPARAMETER TABLE

LEARNING RATE	VALIDATION F1-SCORE
3e-5	0.931
2e-5	0.935

ERROR ANALYSIS

Although DistilBERT performed well, several patterns caused misclassification, some of them are listed in the table below:

ERROR TYPE	EXAMPLE	REASON
Implicit ADR	“Weight loss during therapy”	Requires inference
Hypothetical/uncert-ain Phrasing	“**May** cause allergic symptoms..”	Model expects confirmed statements
Negation Handling	“**No** rash observed”	ignores negation words

CONCLUSION

This project showed that **ADR detection** can be automated using NLP. We tested two types of models - **a basic machine learning model** and a **transformer model.** Some of the insights we could draw by implementing these type models in our project are:

- The basic model worked well but **missed** many ADR sentences. The transformer model **performed much better** and gave **higher accuracy** and better results overall.
- The results prove that transformer models are more suitable for understanding medical sentences.
- This system can help reduce manual work, save time, nd improve patient safety. It can also help in drug monitoring, research, and future healthcare systems.
- With more data and further improvements, this model can become even more reliable and useful in real medical applications.

CONTRIBUTIONS

Anuj (Team leader) :
Evaluated results, generated metrics, created comparative analysis, performed detailed error interpretation.

Abhinav Simha (Data Manager):
Collected, analyzed, cleaned, and split dataset; tokenized input text; documented data characteristics.

Akshay (Logic interpreter):
Implemented baseline and transformer-based models, performed training, tuning, and maintained code base.

Harshith:
Defined problem statement, wrote introduction, and conducted literature review on ADR detection and NLP methods.

Ankit:
Compiled final report, created slides, ensured formatting follows rubric, managed final submission.