International Research Journal of Multidisciplinary Scope (IRJMS), 2024; 5(1): 405-416 IRJMS

Original Article | ISSN (O): 2582-631X DOI: 10.47857/irjms.2024.v05i01.0234

Adverse Drug Reaction Detection from Social Media Review Using BERT Technology

Arijit Dey[1,2]*, Jitendra Nath Shrivastava[2], Chandan Kumar[3]

[1]Department of Computer Applications, B. P. Poddar Institute of Management and Technology, Kolkata, West Bengal, India, [2]Department of Computer Science and Engineering, Invertis University, Bareilly, Uttar Pradesh, India. [3]Young Professional, Ministry of Labour and Employment, Government of India, Noida, India. *Corresponding Author's Email: ad.computerapplication@gmail.com

Abstract

The effects of Adverse Drug Reactions (ADRs) are very harmful to human life; sometimes they cause death. So, the detection of ADRs is very essential. The electronic ADR reporting system has increased, and it is more effective in comparison to the clinical ADR reporting system as patients directly share their opinions on various social media platforms. In the clinical ADR reporting system, it is very difficult to report all the ADRs. ADR extraction from the opinions of patients is very much needed. Advanced Natural Language Processing (NLP) is then essential to apply for the deep learning process. In this article, Bidirectional Encoder Representation from Transformer (BERT) is presented in detection of ADRs from the reviews over the condition on which the drug has been applied. The direction of this work is different from other research work that has been done till now. The condition of the drug to which it has been applied is grouped together to extract the ADRs. This work also helps to develop a drug recommendation system. The f1-score and the accuracy found in the work are very promising, and they outperform the other state-of-the-art machine learning models. This work achieved both a f1-score and accuracy of 0.90 in detecting the ADRs from the analysis of reviews.

Keywords: Adverse drug reaction (ADRs), Natural language processing (NLP), Machine learning, Deep learning, Bidirectional encoder representation from transformer (BERT).

Introduction

Drugs are used to save the lives of human beings and also to improve one's quality of life. Drugs improve the physical and mental fitness status of human beings from several diseases as treatments. Sometimes the use of drugs causes adverse reactions in human beings or patients. The risks of taking medicines are considered during the drug development process. The clinical trials are made to analyze and evaluate the reactions to prevent different adverse drug reactions (ADR). But clinical trials suffer from two major drawbacks to effectively detecting all possible ADRs: i) it takes lots of time; and ii) the limitation of samples. Social media suggests a way to collect opinions from patients directly through online medical forums, such as Twitter, Facebook, blogs, etc. ADRs can be detected by constructing machine learning and deep learning models. Data collected from several social media platforms is analyzed in different aspects, like feature selection, keyword extraction, data balancing, etc. (1). Social media gives people the opportunity to share their opinions in much easier form. The few research centers reported that health topics like disease or treatment have been searched by at least 59% of adult Internet users for seeking information (2). They also reported that 23% of social network users have personal followers of their friends. Health experience (3). The quality of life is degraded due to the adverse effects of drugs, even if they cause death. In a study, it was reported that 3.5% of the patients were hospitalized due to the adverse effect (4). In Europe, it has been predicted that approximately 197,000 deaths happened because of ADR (5). The Food and Drug Administration (FDA) monitors drug safety after releasing the drug (6). The Adverse Event Reporting System (AERS) database is a passive spontaneous reporting system that is suffering from delayed and underreported problems, which leads to an inefficient system (7).

(Received 08th November 2023; Accepted 02nd January 2024; Published 30th January 2024)

To overcome the drawbacks reported earlier to the post-marketing surveillance, social media is used widely now a days (8). Social media is now accessible to almost everybody, and it helps in reporting ADRs in the field of pharmacovigilance. Many researchers have reported the detection of ADRs through various machine learning and deep learning approaches in their studies (9). Support Vector Machine (SVM) (10-14), Naïve Bayse (15), and (16) approaches have been applied by different researchers comparatively in small datasets in the detection of ADRs. Clustering algorithms (17) and (18) are also used in the detection of ADRs. Deep learning, like the Convolutional Neural Network approach (19), (20) has also been used to detect ADRs. Many researchers have applied deep learning techniques to the detection and extraction of ADRs from drugs, which helps the pharmacovigilance organization in the drug development phases.

Literature review

Clinical review suffers from the drawback of not reporting of ADRs in due time. During clinical review, it is very difficult to report all the ADRs. Today, social media like twitter, facebook, blogs, online forums are very much used by the users to share their opinions with others. Social media becomes a way to share the opinions of patients and to report adverse drug effects. Pharmacovigilance companies are also use the benefit of social media during the time of drug discovery. An ensemble model proposed in (21) resulted in a model performance that outperformed all other models submitted in SMM4H 2019. The objective of this work was to create a cutting-edge deep learning model for detecting adverse drug reactions (ADRs) on social media. They have made a contribution by creating a BERT-based deep learning model that combines enhanced domain-specific pre-processing and ensembling to produce cutting-edge outcomes for this ADR classification job. The work also demonstrated the unpredictability that may be anticipated when fine-tuning BERT models on limited datasets and offered some useful advice for how to deal with this variability through averaging. The real-time monitoring of social media platforms for ADRS is the model's ideal use case; however, despite our success in constructing it, they found two key drawbacks in this study. First off, a real-time job may not be well-suited for such a huge ensemble model, and alternate methods, including simpler single-layer BLSTM derived from BERT-based models (22), may be more appropriate. Second, the classification work is simply the first that should be automated; normalization and ADR mention extraction are also necessary. Further study should concentrate on the performance of current models to carry out these downstream tasks, as there is much space for improvement (23). The identification, management, and reporting of ADRS have been covered in their article (24). They have discussed how the use of current technology is altering how ADRS are anticipated, avoided, identified, and handled, as well as how we are constantly working to enhance these procedures. As more phenotypic data can be merged to provide prescribers with patient-specific recommendations, individualized therapy is becoming increasingly feasible. Through the life cycle of a pharmaceutical product, such regulatory science at the national and international levels can assist in achieving a favourable benefit-to-harm ratio. Since avoiding or reducing the risk of ADRS is still a barrier to our routine clinical practice, attaining the greatest results from medicines remains an important priority for individual doctors. A study (25) indicates adopting BERT adjusted with FARM (FARM-BERT) to find ADRS. As a result of its parallelized preprocessing, the proposed model, FARM-BERT, is computationally faster than normal BERT and is hence appropriate for usage in production settings. An all-encompassing method for diagnosing ADRS is provided using multitask learning. The downstream task of recognizing ADRS is modelled by BERT, which is first tuned with FARM using pretraining on the BBC news corpus. Results from experiments on the Twitter, PubMed, and TwiMed datasets are compared with those using SVM, MLP CNN, LSTM, and standard BERT baseline models. The outcomes are also contrasted with other cutting-edge research. On Twitter, PubMed, TwiMed-Twitter, and TwiMed-PubMed datasets. However, BERT pretrained on the BBC corpus is employed as the language model in the downstream job of the biomedical domain, despite the fact that the results of the proposed approach are extremely encouraging. The linguistic specifics of the domain

of the downstream task cannot be accurately represented by using a language model from a different domain. By pretraining the BERT language model on the biomedical text and modifying it for the downstream job of ADR detection, the suggested approach can therefore be further improved. By looking at the impact of pretraining BERT upon this biological text and then refining it with FARM to detect ADRS, we hope to overcome this constraint in the future.

Proposed model

This article proposes a novel approach to find the adverse reactions of a drug or drugs based on the condition to which they have been applied and also being able to recommend the best-suited drug for the condition. The proposed model consists of five major steps: i) data collection; ii) pre-processing; iii) grouping; iv) tokenization; and v) passing to the BERT model for pre-training and classification. The following figure 1 illustrates the proposed model.

Data collection

Collect raw data from the data source (26) that is described in this section. The patient's opinion is shared through several online forums, twitter etc. A review dataset has been found where the condition on which the drug has been applied has been shared by several patients. This dataset is a review dataset in which drug id, condition, review, ratings, useful counts, dates, and other attributes are described. The opinions are shared directly by patients in this dataset. The approach proposed in this article is different from other works done on this dataset. The data is publicly available on GitHub, allowing users to choose their research use without any identifiable information. This ensures privacy and acknowledges and addresses biases. The dataset is cited from its download location, and the authors are cautious about generalizations that may perpetuate biases. They also strive to diversify datasets and address limitations while being cautious about making generalizations that may perpetuate biases.

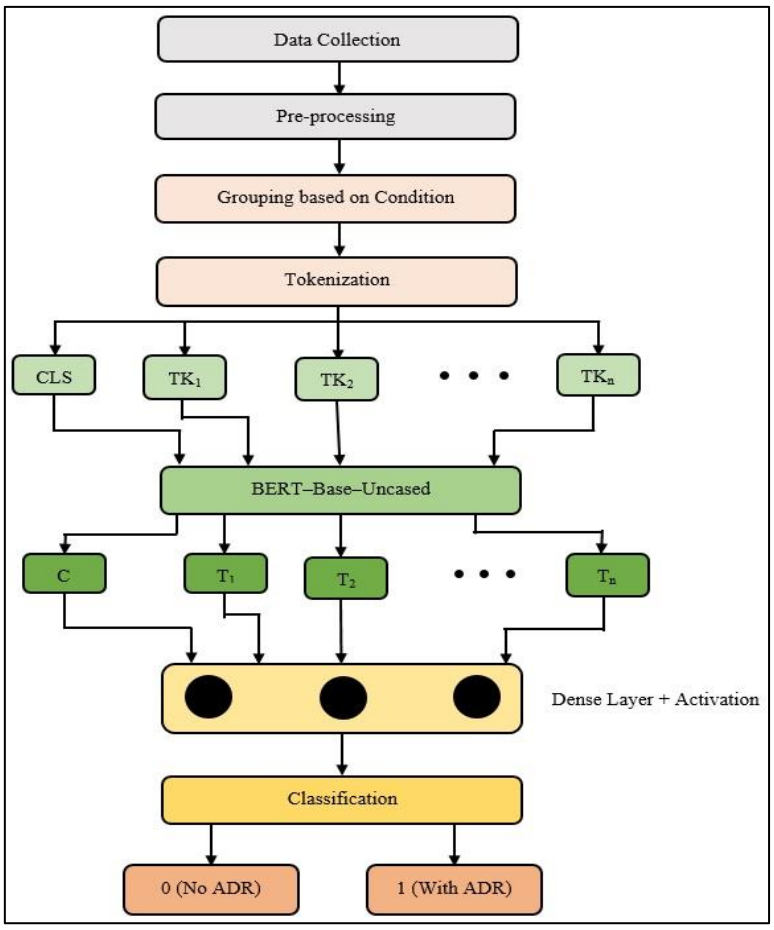


Figure 1: Steps of Proposed Model

Pre-processing

This section is a mandatory part of working on this dataset. The review comments contain different types of punctuation, numbers in between letters, as well as various mentions as the patients directly share their views on any drug in this dataset. To remove all the above mentioned is very much necessary to apply natural language processing to the review comments. This article proposes a model that detects the presence or absence of ADRs in drugs. The values range from 1 to 10, given as ratings in this dataset. So, this article adds an extra attribute called'sentiment' and puts values as '0' for non-ADR review comments, which have ratings ranging from 9 to 10, and puts values '1' in those cells for review comments that contain ADR, where the ratings lie between 1 and 8. Basically, this article proposes that the absence of ADR is represented by higher-order ratings, and lower-order ratings represent the presence of ADR. To ensure accurate analysis, it is crucial to identify and handle missing values, outliers, errors, and inconsistencies in the dataset. This can be done by removing rows with missing values using statistical methods or advanced imputation techniques. Correct errors and inconsistencies in the dataset, such as fixing typos or resolving inconsistencies in categorical values. Transform the data to meet analysis assumptions or make it more suitable for modeling. Identify and handle duplicate records to avoid biased results. Convert categorical variables into a suitable format for analysis using encoding methods.

Grouping

In this section, the article proposes a new approach before analyzing various review comments on drugs by patients. In the study, it has been noticed that a drug has been used in different conditions, or in the same condition, multiple drugs have been used. A grouping mechanism has been used before applying any machine learning or deep learning technique. All the drugs are grouped together based on the condition they have been applied for. The article grouped together the drugs along with all the review comments of patients based on the most common conditions shown in figure 2. The article also found those drugs which have been used for many conditions, as shown in figure 3. Figure 4 shows the most drugs available per condition per patient. The first 20 samples have taken in the three figures (figure 2, figure 3, and figure 4) for better visualization.

Tokenization

The process of fragmenting raw text into small chunks is known as tokenization. It is important as the raw text or sentences cannot be read by any machine learning or deep learning models. Tokenization helps in the development of models and has an effect on their performance. Tokenization can be word tokenization or sentence tokenization. In this article, the Keras model has been used to achieve the desired result.

BERT model

This section describes that pre-training BERT and fine-tuning BERT are the two-step frameworks used in this proposed model. Pre-training BERT has been used on unlabeled data over several pre-training tasks during the pre-training session. Then pre-trained parameters have been set to the BERT model and fine-tuned to all the necessary parameters from the set of labeled data (27). It affects the performance of the algorithm in terms of learning speed (28). One of the major reasons for the wise performance of the BERT algorithm is that it learns the linguistic patterns and transforms the facts during the pre-training of the model. The model receives knowledge after training and is built by fine-tuning the parameters. An attention mechanism is used to learn the contextual relationship between the words. BERT technology uses this attention mechanism to find the relationship between the words in reviews given by the patients (29). In this work, the drugs are grouped together based on the condition, and the attention mechanism of BERT technology is used to extract the ADRs from the drugs for which it has been used. This article uses the $BERT_{BASE}$ model, where 12 encoders are used for the vectorization process. The following figure 5 shows the BERT architecture. The proposed model is a text categorization since it can only determine whether a medication contains adverse drug reactions (ADRs) or not. The suggested model uses WordPiece tokenization since it is based on the bert-based-uncased model. Typically, input sequences have a unique [CLS] token at the beginning and [SEP] tokens in between sentences. The tuning parameter for an optimization approach that

determines the step size at each iteration while moving toward a minimum of a loss function is called the learning rate, and it has been set at 5e-5. A neural network's single forward and backward traversal of an entire dataset is known as an epoch. Seven is a representation of how frequently the algorithms examine all of the data. The batch size for a forward or backward pass is set to six training instances overall. In Google Colab (normal version), the proposed approach is employed. As the model is trained on a large dataset, there are a number of hardware limitations. The hyperparameters have been adjusted to yield a more precise outcome. The description of the details of the proposed BERT model is shown in Table 1.

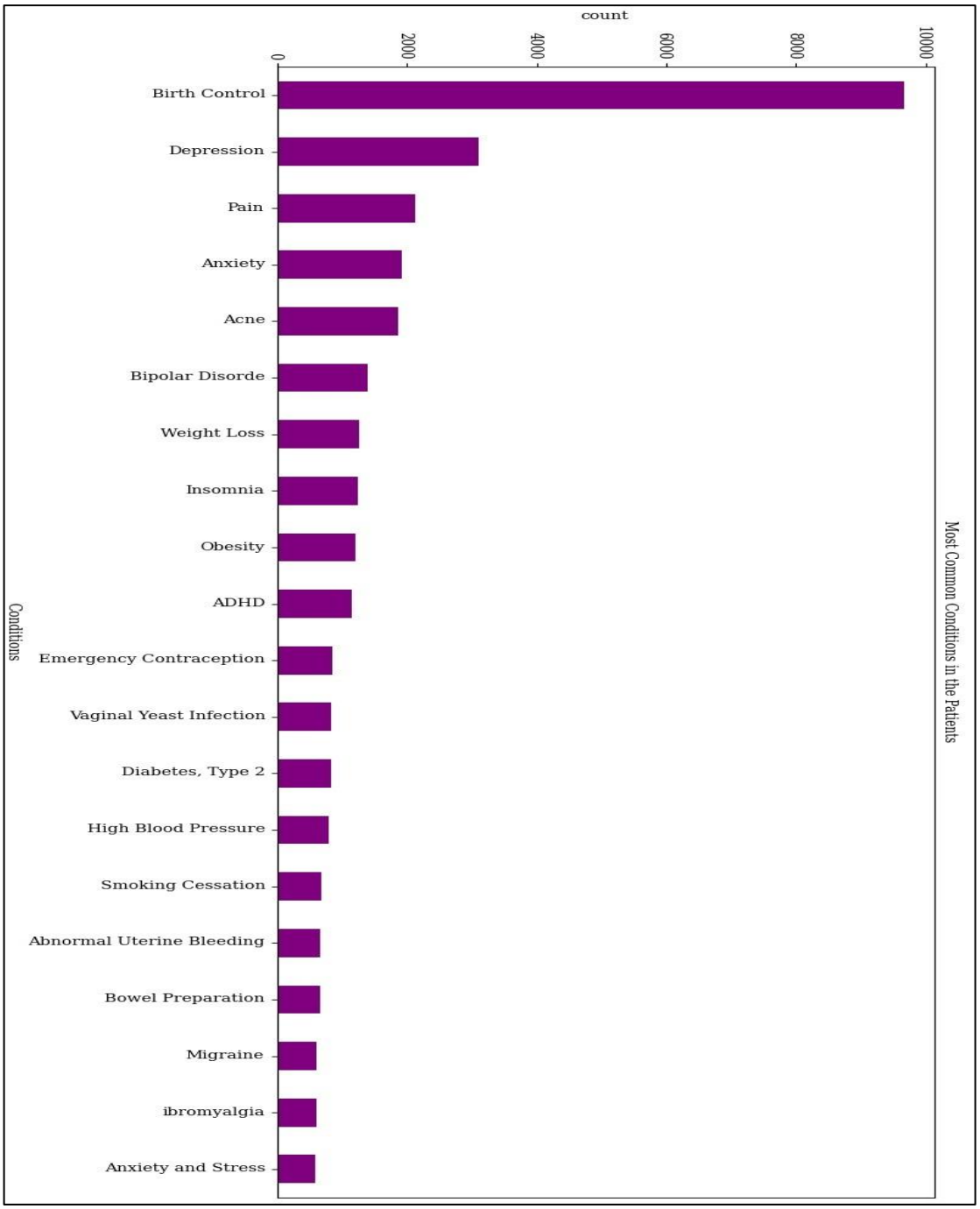


Figure 2: Most Common Condition in Patient

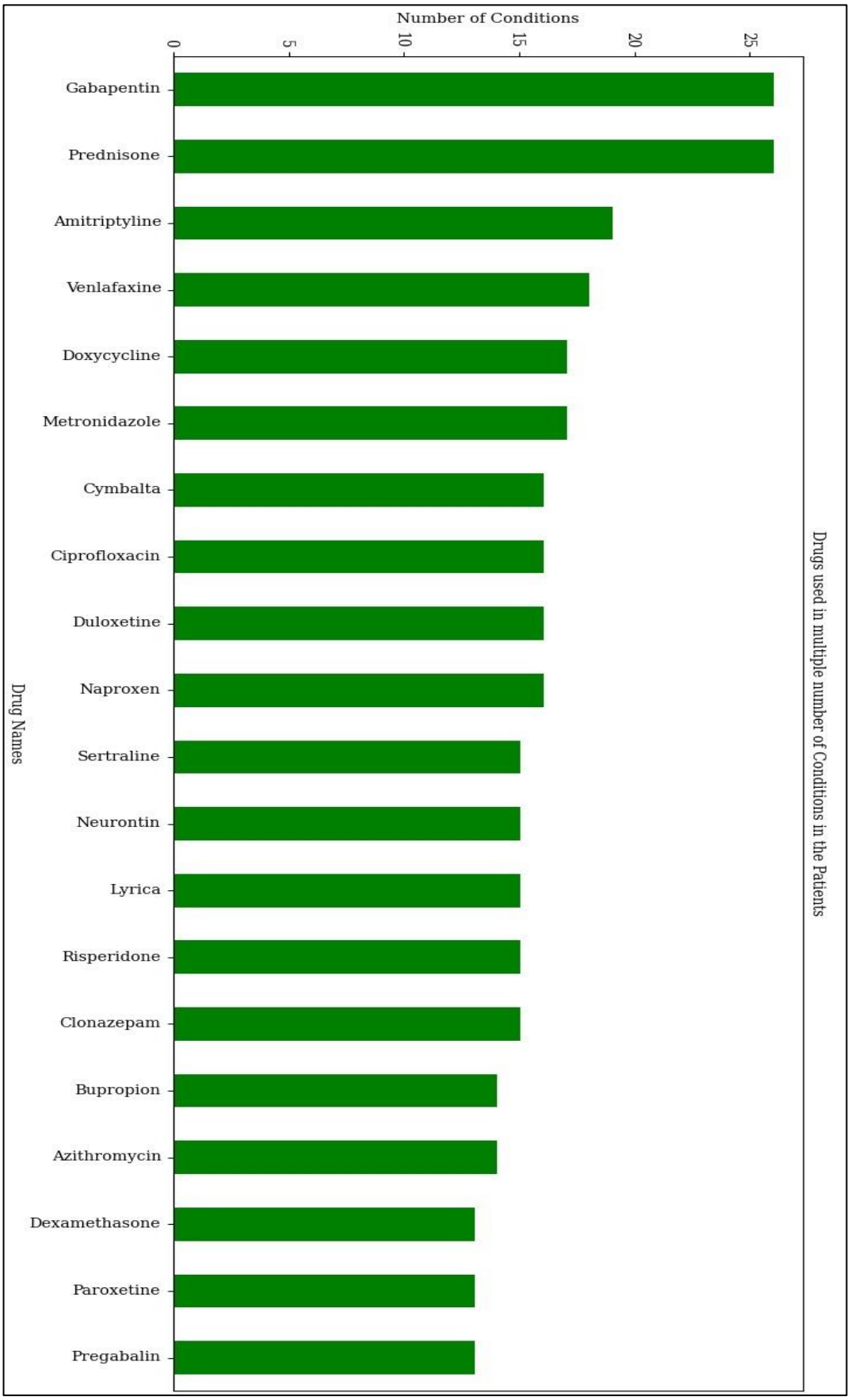


Figure 3: Drugs Used Many Conditions

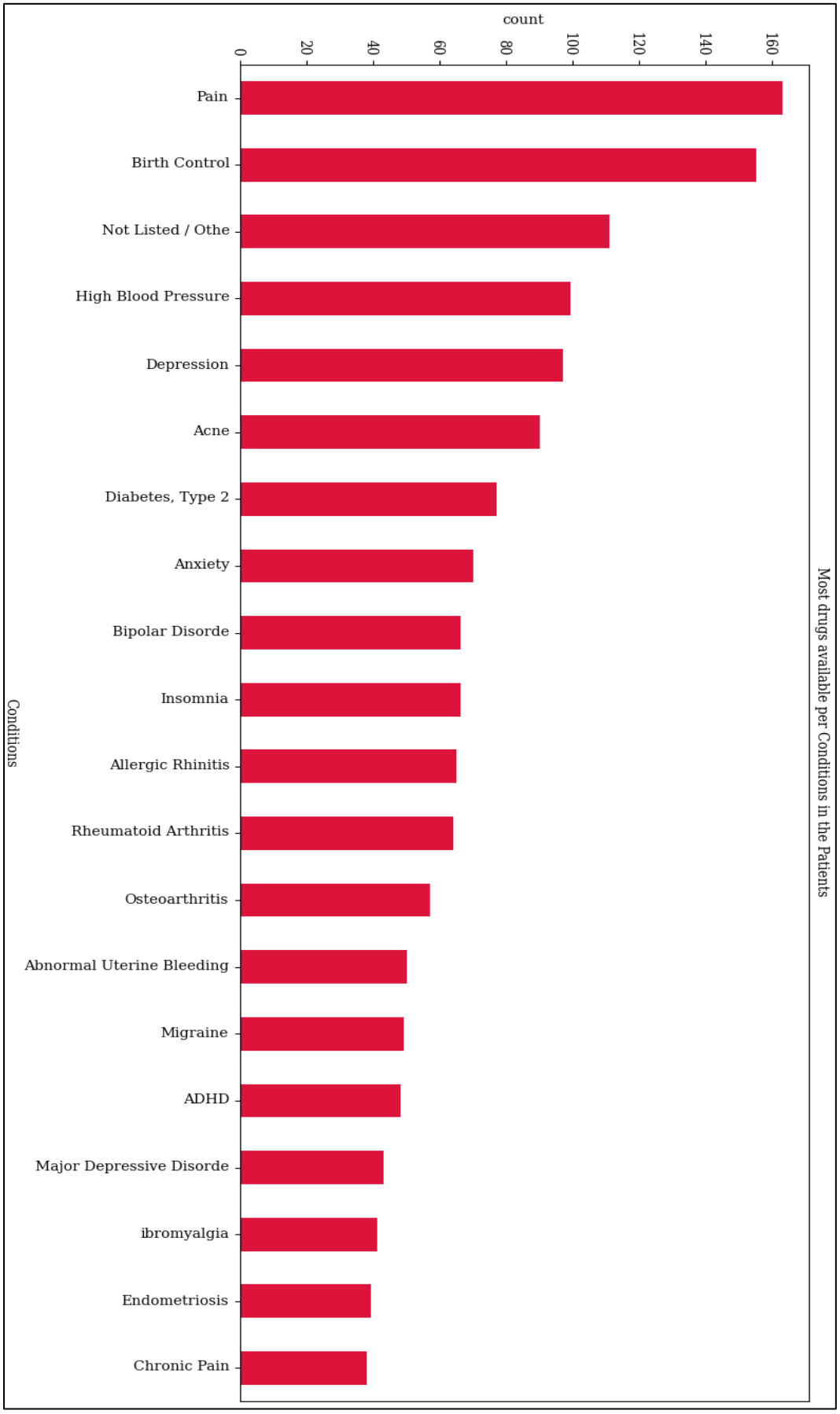


Figure 4: Most Drugs Available per Condition

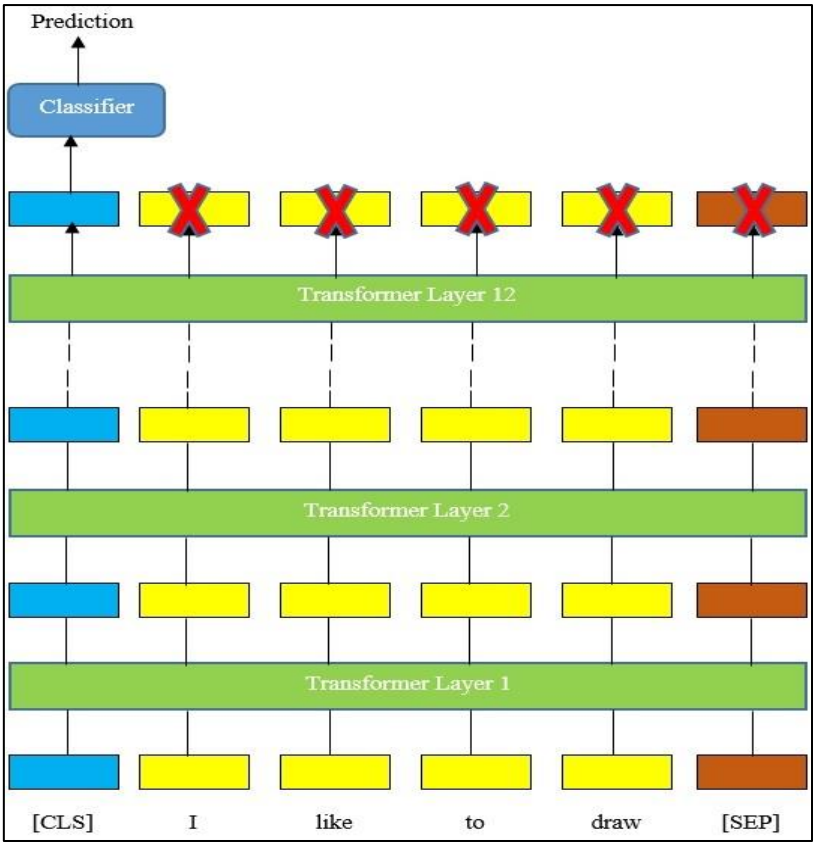


Figure 5: BERT-Architecture

Table 1: Parameters selected for proposed BERT model

Parameters	Data / Value
vocab_size	30522
hidden_size	768
num_hidden_layers	12
num_attention_heads	12
hidden_act	'gelu'
max_position_embedding	512
type_vacab_size	2
epochs	7
learning_rate	5e-5

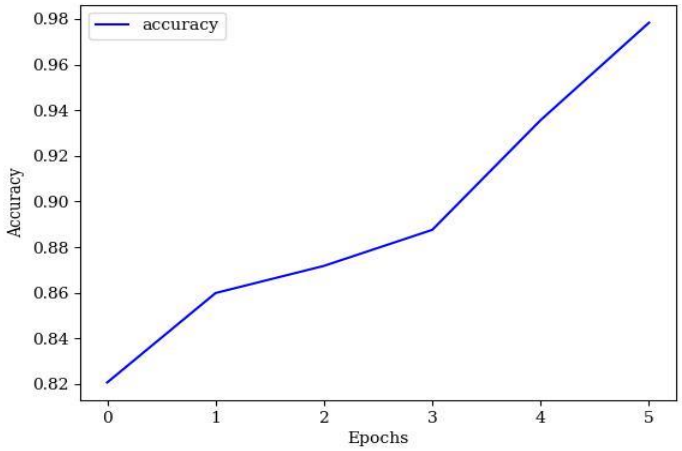
Results

The proposed model outperforms over the state-of-the-art models discussed earlier. The performance measurement of the proposed model has been achieved using some of the metrics. The performance measurement metrics are accuracy, precision, recall, and f1-score. The reasoning behind choosing these evaluation matrices for the proposed BERT model is that because the classes are imbalanced, accuracy—which serves as a general indicator of the model's correctness—may be misleading. When the cost of false positives is large, accuracy is crucial. It gauges how accurate positive forecasts are. When fetch has a high rate of false negatives, recall is essential. It assesses how well the suggested model can account for every real positive case. The F1 score offers a single metric that accounts for both precision and recall stability, which is crucial in certain situations.

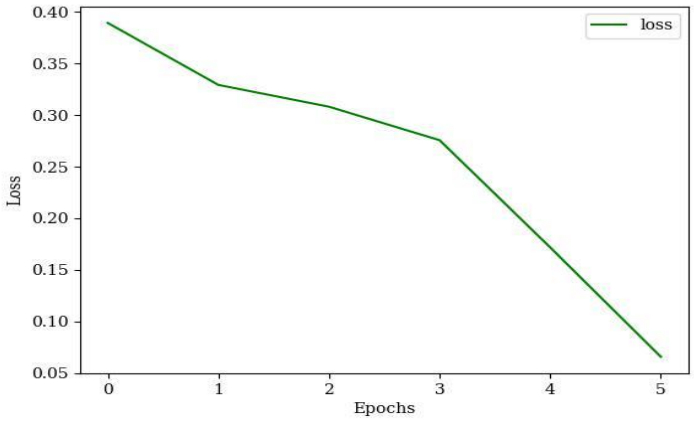
The pre-training BERT model has been processed on the training dataset. The figure 6(a) shows the accuracy, and figure 6(b) shows the loss over the

training data. The proposed model has been trained with a one-cycle policy, considering a maximum learning rate 0.00005. The loss became 0.0659 and achieved accuracy 0.9784 after the 6 epochs over training dataset.

The performance of the proposed model has been evaluated by the following evaluation criteria shown in equations 1, 2, 3, and 4. The achieving accuracy over testing data is shown in table 2. The comparison of the proposed model with other state-of-the-art models is shown in figure 7.



(a)



(b)

Figure 6: (a) Training accuracy, (b) Loss after 6 epochs

$$accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad [1]$$

$$precision = \frac{TP}{TP + FP} \quad [2]$$

$$recall = \frac{TP}{TP + FN} \quad [3]$$

$$f1 - score = \frac{2 \times precision \times recall}{precision + recall} \quad [4]$$

Table 2: Accuracy table

Values	Precision	Recall	F1-score
0	0.89	0.91	0.90
1	0.91	0.89	0.90
accuracy			0.90

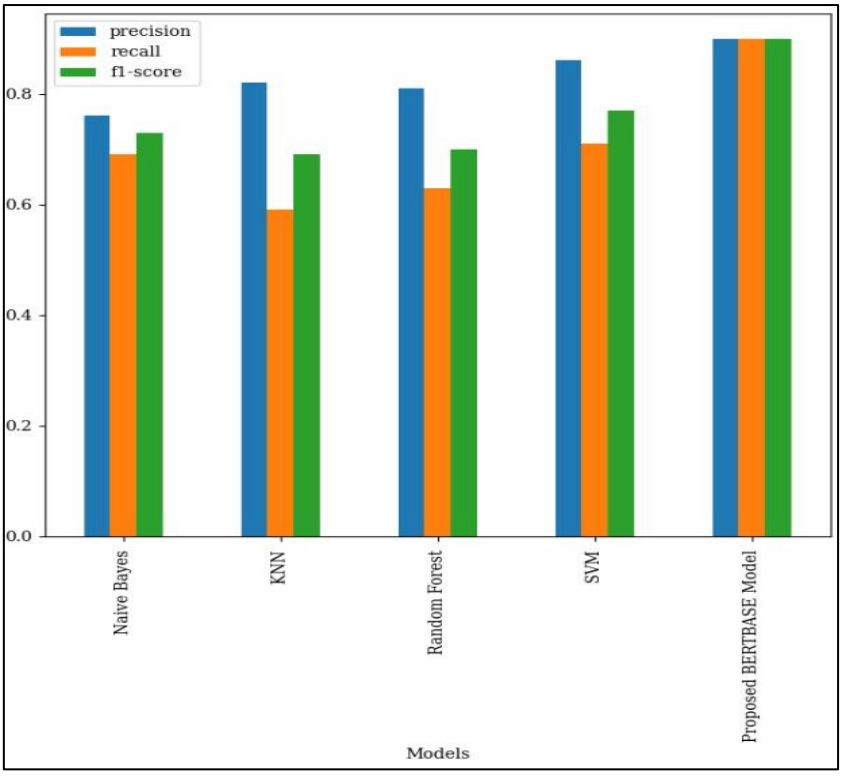


Figure 7: Comparison of proposed model with other models

Conclusion

The basic approach of various machine learning models have been designed, but these algorithms have failed to reach high-performing models in the detection of ADRs. So deep learning models have been introduced earlier by several researchers. BERT architecture has also been proposed in sentiment analysis. In this article, BERT architecture has been used in the detection of ADRs of drug or drugs that may be used in several conditions. It is also being seen that some of the drug has been applied to different conditions. This paper grouped together the conditions so that all the drugs and their corresponding reviews can be found in a group. After analyzing the reviews, it is determined whether the drug contains any ADRs or not. The experimental results discussed above show that the proposed model provides a promising result in condition-based ADRs for the detection of drugs. This work helps the pharmacovigilance company with drug discovery. In this work, the other state-of-the-art machine learning-models (Naive Bayse, KNN, Random Forests, and Support Vector Machines) are compared with the proposed $BERT_{BASE}$ model and show a significant improvement. The attention mechanism of Transformer is used in the BERT model and learns the contextual significance of the words used in the text. As this work is implemented from text reviews of patients', it is very important to learn the contextual relationship between the words. This work learns the contextual relationship between the words during pre-training and the fine-tune of the $BERT_{BASE}$ model to detect the ADRs as a concept of knowledge transfer. The proposed BERT model is pre-trained on large, diverse datasets, which might not be specific to the biomedical or pharmacological domain. This can limit their ability to understand and extract domain-specific information related to adverse drug reactions accurately. The proposed model is complex and often lacks interpretability, making it challenging to understand in medical applications. This is especially important in the dynamic healthcare domain, where new terminologies and concepts are frequently introduced. Without continuous updates, the proposed models may struggle to adapt to the evolving medical language.

Abbreviations

Nil

Acknowledgement
Nil

Authors contribution
All authors contributed equally in this proposed work.

Conflict of interest
The authors declare no conflict of interest.

Ethics approval
Not applicable

Funding
There is no fund applicable for this study.

References
1. Huang JY, Lee WP, Lee KD. Predicting adverse drug reactions from social media posts: data balance, feature selection and deep learning. In Healthcare (MDPI). 2022 ;10(4):618
2. Odeh F, Taweel A. A deep learning approach to extracting adverse drug reactions. In: 2019 IEEE/ACS 16th International Conference on Computer Systems and Applications (AICCSA). 2019 Nov 3; 2019. p. 1-6.
3. Fox S. The social life of health information. 2011.
4. Bouvy JC, De Bruin ML, Koopmanschap MA. Epidemiology of adverse drug reactions in Europe: a review of recent observational studies. Drug safety. 2015 May;38:437-53.
5. Commission E. Proposal for a regulation amending, as regards pharmacovigilance of medicinal products for human use. Regulation (EC) No 726/2004. Impact assessment. 2008.
6. World Health Organization. The importance of pharmacovigilance. 2002.
7. Ahmad SR. Adverse drug event monitoring at the Food and Drug Administration: your report can make a difference. Journal of general internal medicine. 2003 Jan;18:57-60.
8. Dey A, Shrivastava JN, Kumar C, Chandra S. Adverse Drug Reactions Extraction from Social Media: A Systematic Review. Grenze International Journal of Engineering & Technology (GIJET). 2022 Jan 1;8(1): 91 – 98.
9. Aroyehun ST, Gelbukh A. Detection of adverse drug reaction in tweets using a combination of heterogeneous word embeddings. In Proceedings of the fourth social media mining for health applications (# SMM4H) workshop & shared task. 2019 Aug; 2019; p.133-135.
10. Bollegala D, Maskell S, Sloane R, Hajne J, Pirmohamed M. Causality patterns for detecting adverse drug reactions from social media: text mining approach. JMIR public health and surveillance. 2018 May 9;4(2):e8214.
11. Chowdhury S, Zhang C, Yu PS. Multi-task pharmacovigilance mining from social media posts. In Proceedings of the 2018 World Wide Web Conference. 2018 Apr 23; 2018. p. 117-126.
12. Islam T, Hussain N, Islam S, Chakrabarty A. Detecting adverse drug reaction with data mining and predicting its severity with machine learning. In2018 IEEE region 10 humanitarian technology conference (R10-HTC). 2018 Dec 6; 2018. p. 1-5.
13. Miranda DS. Automated detection of adverse drug reactions in the biomedical literature using convolutional neural networks and biomedical word embeddings. arXiv preprint 2018 Apr 24; 2018. arXiv:1804.09148.
14. Ding P, Zhou X, Zhang X, Wang J, Lei Z. An attentive neural sequence labeling model for adverse drug reactions mentions extraction. IEEE Access. 2018 Nov 20; 6:73305-15.
15. Li Y, Jimeno Yepes A, Xiao C. Combining social media and FDA adverse event reporting system to detect adverse drug reactions. Drug safety. 2020 Sep; 43:893-903.
16. Yousef RN, Tiun S, Omar N. Extended trigger terms for extracting adverse drug reactions in social media texts. Journal of Computer Science. 2019;15(6):873-9.
17. Laksito AD, Sismoro H, Rahmawati F, Yusa M. A comparison study of search strategy on collecting Twitter data for drug adverse reaction. In 2018 International Seminar on Application for Technology of Information and Communication. 2018 Sep 21; 2018. p. 356-360.
18. Yang CC, Zhao M. Determining associations with word embedding in heterogeneous network for detecting off-label drug uses. In 2017 IEEE International Conference on Healthcare Informatics (ICHI). 2017 Aug 23; 2017. p. 496-501.
19. Min Z. Drugs reviews sentiment analysis using weakly supervised model. In 2019 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA). 2019 Mar 29; 2019. p. 332-336.
20. Zhang M, Geng G. Adverse drug event detection using a weakly supervised convolutional neural network and recurrent neural network model. Information. 2019 Sep 4;10(9):276.
21. Breden A, Moore L. Detecting adverse drug reactions from twitter through domain-specific preprocessing and bert ensembling. arXiv preprint. 2020 May 11; 2020. arXiv:2005.06634.
22. Chen S, Huang Y, Huang X, Qin H, Yan J, Tang B. HITSZ-ICRC: a report for SMM4H shared task 2019-automatic classification and extraction of adverse effect mentions in tweets. In Proceedings of the fourth social media mining for health applications (# SMM4H) workshop & shared task. 2019 Aug; 2019 p. 47-51.
23. Weissenbacher D, Sarker A, Magge A, Daughton A, O'Connor K, Paul M, Gonzalez G. Overview of the fourth social media mining for health (SMM4H) shared tasks at ACL 2019. In Proceedings of the fourth social media mining for health applications (# SMM4H) workshop & shared task. 2019 Aug; 2019. p. 21-30.
24. Coleman JJ, Pontefract SK. Adverse drug reactions. Clinical Medicine. 2016 Oct; 16(5):481.
25. Hussain S, Afzal H, Saeed R, Iltaf N, Umair MY. Pharmacovigilance with transformers: A framework to detect adverse drug reactions using BERT fine-tuned with FARM. Computational and

Mathematical Methods in Medicine. 2021 Aug 13; 2021.
26. MichaelAllen1966/1804_python_healthcare; Dec 11, 2020. Available from: https://github.com/MichaelAllen1966/1804_python_healthcare/tree/master/data
27. Devlin J, Chang MW, Lee K, Toutanova K. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint. 2018 Oct 11; 2018. arXiv:1810.04805.
28. Harnoune A, Rhanoui M, Mikram M, Yousfi S, Elkaimbillah Z, El Asri B. BERT based clinical knowledge extraction for biomedical knowledge graph construction and analysis. Computer Methods and Programs in Biomedicine Update. 2021 Jan 1; 1:100042.
29. Dey A, Shrivastava JN, Kumar C. Transformer Based Knowledge Graph Construction in Adverse Drug Reactions Prediction from Social Media. 2022 Oct; 70(10): 402-407.