An Ensemble Model for Detection of Adverse Drug Reactions

Ahmed A. Nafea[1,*], Mustafa S. Ibrahim[2], Abdulrahman A. Mukhlif[3], Mohammed M. AL-Ani[4] and Nazlia Omar[4]

[1]Department of Artificial Intelligence, College of Computer Science and IT, University of Anbar 31001, Ramadi, Anbar, Iraq

[2]Department of Computer Science, University of Anbar 31001, Ramadi, Anbar, Iraq

[3]Registration and Students Affairs, University Headquarter, University of Anbar 31001, Ramadi, Anbar, Iraq

[4]Center for Artificial Intelligence Technology (CAIT), Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia (UKM), Bangi, Selangor, Malaysia

Abstract—**The detection of adverse drug reactions (ADRs) plays a necessary role in comprehending the safety and benefit profiles of medicines. Although spontaneous reporting stays the standard approach for ADR documents, it suffers from significant under-reporting rates and limitations in terms of treatment inspection. This study proposes an ensemble model that combines decision trees, support vector machines, random forests, and adaptive boosting (ADA-boost) to improve ADR detection. The experimental evaluation applied the benchmark data set and many preprocessing techniques such as tokenization, stop-word removal, stemming, and utilization of Point-wise Mutual Information. In addition, two-term representations, namely, term frequency-inverse document frequency and term frequency, are utilized. The proposed ensemble model achieves an F-measure of 89% on the dataset. The proposed ensemble model shows its ability in detecting ADR to be a favored option in achieving both accuracy and clarity.**

***Index Terms*—Adverse drug reactions, Classification, Ensemble Model, Machine Learning, Point-wise Mutual Information.**

1. Introduction

Adverse drug reactions (ADRs) are a harmful side effect that happens when patients are taking drugs. These effects can be severe and may arise either from the pharmacological properties of the drug, interactions with other drugs, or existing medical conditions (Edwards and Aronson, 2000; Kiritchenko; Zhu and Mohammad, 2014). ADRs establish a significant public health concern, with approximately 2.2 million serious cases occurring annually in the United States alone. Identifying and monitoring ADRs are vital to ensuring patient safety and appropriate medication usage (Yadesa et al., 2021; Ebrahimi et al., 2016). Early detection of ADRs is necessary in avoiding consequences and improving patient effects. The development of accurate and efficient is importance methods for detecting ADR (Kiritchenko et al., 2018).

Machine learning is an automatic detection of ADR from unstructured clinical text and has been greatly enabled by latest developments in natural language processing (NLP) and ML. This improvement is useful into the large data within health-care systems using these methods focused on data that show significant in ADR detection and elevating patient safety (Sørup et al., 2020). It is a must to implement challenges, like accuracy of ADR detection, to fully understand the advantages presented by these methods.

This study focuses on detecting ADR by the proposed of an ensemble model. This study utilized random forest (RF), support vector machines (SVM), decision trees (DT), and ADA-boost algorithms, with the aims to improve the accuracy and efficiency of ADR detection. This study utilized dataset from a benchmark collection by (Yates and Goharian, 2013), and this study by Yousef et al. (2019) has been extended this dataset by combining supplementary meaningful attributes.

The organization of this study as follows: Firstly, a comprehensive review of previous research on ADR detection using ML is presented. The second section describes the methodology applied in this study which includes the pre-processing, feature engineering, and model training. The results of the study, including performance metrics of the ensemble model utilizing RF, SVM, DT, and ADA-boost algorithms, are then presented and compared to different approaches. Finally, the effects of the findings are discussed and potential ways for future research on ADR detection using deep learning techniques.

ARO-The Scientific Journal of Koya University
Vol. XII, No. 1 (2024), Article ID: ARO.11403. 7 pages
DOI: 10.14500/aro.11403
Received: 14 September 2023; Accepted: 25 January 2024
Regular research paper: Published: 20 February 2024
Corresponding author's e-mail: ahmed.a.n@uoanbar.edu.iq

II. Related Work

Wang et al. (2019) proposed a deep neural network (DNN) to develop model reserved for the automated detection of ADR, utilizing chemical, biological, and biomedical information pertaining to drugs. This model was designed with two primary objectives. The first is the sensitivity of potential ADR related to established drugs. The second one prediction of possible ADRs linked to novel drugs. The model combined word-embedding techniques to efficiently get complex drug relationships present in extensive biomedical literature to account for new drugs, to another place from the dataset, a mapping function was created.

In this study by Zhang et al. (2020), the authors proposed a Gated Iterative Capsule Network (GICN) was established as an modern model for detecting ADR. This model combined character embedding to proficiently operate abbreviations and typographical errors. Given the multi-word nature of ADR, the model utilized CNN to cover comprehensive phrase-related information to extract deep semantic degrees, a case network featuring a gated iteration unit was proposed. This mechanism helped the classified grouping from lower level to higher-level cases while retaining related information. Experimental findings show that the GICN model displayed enhanced performance in ADR prediction from social media text compared to other current approaches. This model's efficacy contributes significantly to the improvement of ADR detection within the realm of text analysis.

This research by Yousef et al. (2020) was focused on the problems around the identification of ADR within the large area of medical information accessible via social networks. The previse studies methods commonly depend on medical dictionaries for ADR extraction often utilized trigger terms or text extension techniques. These techniques have a limitation regarding the treatment of abbreviations and effect on whole text contexts. This study proposed a lexicon alternative approach relate on the replacement of individual terms rather than entire sentences to overcome these limits. There are different previous works; this approach combined a medically pre-trained word embedding model to support in another task. Experimental evaluations including medical review benchmark datasets and three classifiers SVM, LR, and NB showed developments in classification accuracy through the proposed lexicon replacement method. This contribution highlights the advantage of the created approach compared to normal techniques in the field of ADR extraction from social network data.

This research by Li et al. (2020) focused on difficult challenged task of detection ADR. This study proposed models for ADR detection showed limitations including small-scale benchmark measures and the necessary for additional manually explained amounts or co-training with entity-mentioning extraction tasks, which might introduce noise or escalate explanation pains to address these challenges, this researcher focuses on ADR detection as a text classification task and introduced an adversarial transfer learning framework. This approach controlled a source quantity to reinforce performance within limited training cases in smaller target measures. Adversarial learning was deployed to avoid the advance of corpus-specific features into the shared area, ensuring the effective utilization of diverse corpora. The experiential results among three benchmark corpora supported the advantage of the proposed method, particularly in the context of small-scale corpora, outperforming existing state-of-the-art strategies. This study proposed transfer learning and adversarial mechanisms for improving ADR detection.

This paper by Zhang et al. (2021) focuses on the complex task of detecting ADR entities within text, uniquely within the field of social media data. While social media offers real-time and dynamically evolving drug reaction information, the lack of explained social media data has presented a challenge for research in this area. In addition, the informal and informal expressions prevalent in social media posts introduce significant problems for ADR-named entity recognition (NER). To improve these difficulties, the study proposed an adversarial transfer learning architecture for ADR NER. This architecture benefits from on biomedical domain information derived from PubMed to enhance performance within Twitter data. This proposed approach achieved state-of-the-art performance without depending on manually engineered features, yielding a notable F1 score of 68.58% on Twitter ADR data. This research highlights the might of adversarial transfer learning in addressing the challenges unique to ADR identification in the context of social media data.

Chen et al. (2021) proposed focus on challenged of detecting ADR events using social media data, particularly from platforms like Twitter. Conventional post-marketing surveillance systems, trusting on natural reports, are susceptible to underreporting issues, needing alternative data sources. However, research within this domain works with limitations stemming from limited explained datasets, which can delay the efficacy of deep learning models that often trust on large training samples. In response, the study introduced two regularization techniques at the representation level graph embedding-based data augmentation and adversarial training. These techniques aimed to amplify the performance of ADR within the constraints of data lack. The study analyzed and deliberated upon the applicability of these techniques through rigorous experiments. This study proposed an adverse drug event detection framework merging these regularization methods with a convolutional neural network to extract the full advantages. This research significantly contributes to filtering the detection of adverse events in the field of social media data and presents innovative regularization strategies to implement challenges related with limited explained datasets.

In Nafea, Omar and AL-Ani (2021), this study proposed LSA to detecting ADR from social networks, where individuals articulate their perspectives on medications. The previous studies predominantly leaned on trigger terms for ADR detection, this method needed regular updates to hold novel side effects and pertinent medical entities. The feature space built only on trigger terms lacked latent semantic comprehension. To surmount these limitations, the study proposed a semantic approach rooted in LSA to enhance

ADR detection. Experimental studies concerned a benchmark dataset and encompassed preprocessing operations such as stopping word removal, tokenization, and stemming. Three classifiers (SVM, NB, and LR) were trained on the proposed LSA, useful two document representations as TF and TFIDF. The results underscored the superiority of the proposed LSA methodology over the baseline extended trigger term approach, achieving an F-measure of 82% on the dataset. This elevation highlights the effectiveness of LSA in identifying accurate semantic similarities, transcending the utility of predefined trigger term lists. The study explains the relevance of combining semantic insights for ADR detection, thereby contributing to advancements within this domain.

In Nafea, Omar and Al-qfail (2023), this study focuses on the extraction of ADR from user-generated comments and reviews. While previous research mostly focused on machine learning techniques for ADR detection and applied noted medical review data for training classification models, the domain still challenges with relating to detection accuracy. To solve these complexities, this study introduced a composite approach involving LSA and ANN classifiers for the accurate ADR detection. Experimental findings supported the efficacy of merging LSA in tandem with ANN for accurate ADR extraction. This study has the potential to refine ADR detection methodologies and highlights the significance of combining LSA and ANN classifiers to realize accurate ADR extraction.

The key limitation of previous studies depends just on word embeddings that focused on term sequences and required pre-training the model specifically using those embeddings and they used only one model to detect ADR; there is still area for improving the accuracy of ADR detection. This study aims to address these challenges by proposed an ensemble model to enhance ADR detection performance. The proposed ensemble model approach combines multiple individual models as RF, SVM, DT, and ADA-boost to achieve more accurate predictions compared to using a single model. By utilizing and combining different models, the ensemble can capture a wider range of information leading to improved accuracy of ADR detection.

III. Research Methodology

The method of this study contains five stages, as shown in Fig. 1. The first step shows the explanation of drug reviews through the use of a dataset from Yates and Goharian (2013) benchmark datasets. This data set has been enhanced with additional valuable data fields by Yousef, Tiun, and Omar (2019). Following that, several preprocessing tasks include stemming, stop word removal, and tokenization. Then, the terms within the drug reviews are represented in a vector space using term frequency-inverse document frequency (TF-IDF) and TF. After that, PMI is applied. Finally, the classification process uses an ensemble model that combines RF, SVM, DT, and ADA-boost algorithms. The following section shows a more comprehensive and detailed elucidation of the research methodology.

A. Dataset

This study used dataset created by Yates and Goharian (2013), which was enhanced with additional informative data columns from Yousef et al. (2019). The dataset utilized in this research consisted of 2500 reviews, out of which 246 were labeled documents. Every document contained one or more sentences. The texts extracted from Twitter contained a total of 945 sentences. Among all the texts, there were 982 ADR identified. The documents were written in English as shown in Table I; the dataset details in Table II show a sample of dataset. Data from three prominent platforms specializing in drug reviews on social media, namely, askapatient.com,

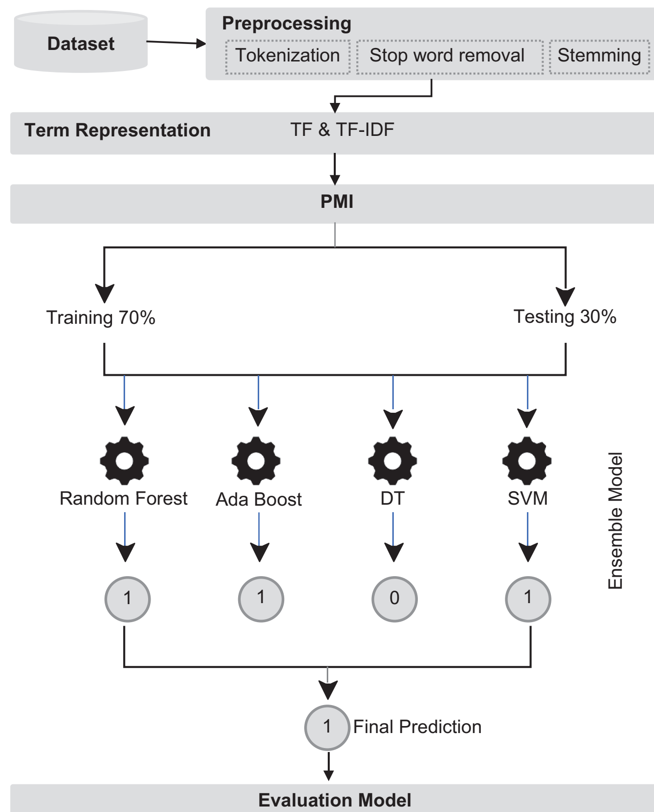

Fig. 1. Research methodology.

TABLE I
Dataset Details

Attribute	Total
Number of ADR	982
Number of Reviews	2500 (labeled 246)
Number of Sentences	944

ADR: Adverse drug reactions

TABLE II
Sample of the Dataset

Doc	Sen	Class	Review	ADR
1	1	0	I had been on Tamoxofin for 5 years.	[]
1	2	1	I still have the neuropathy, but I can deal with that	['neuropathy']
2	1	1	The night sweats were the worst!	['sweats']
3	1	0	I have been taking Femara for 5 months. []	[]
4	1	1	Lower back pain.	['pain']
4	2	0	Vaginal dryness is the other symptom I have.	[]

drugratingz.com, and drugs.com, were gathered for the review analysis. There are two class types the first one ADR is 1 and the second non-ADR is 0.

B. Preprocessing

Pre-processing plays an essential role in NLP, as it shows the cleansing and transformation of unstructured text data into a suitable format for analysis and modeling using machine learning algorithms. This study applies three preprocessing techniques as a stop word removal, stemming, and tokenization.

- Tokenization: It is an initial step in preprocessing, where the text is divided into individual words or tokens (Oyebode and Orji, 2023).
- Stop words: It is commonly used words that have limited meaning. There is an example of these words including “the,” “and” “of,” and “to.” the removal of these words reduces the dimensionality of text data, leading to improved performance of ML models by reducing noise and improving their ability to extract valuable information (McMaster et al., 2023).
- Stemming: It is used to cause the reduction of a word to its fundamental showed by the move of words like “running” to “run.” This approach helps the reduction of the number of unique words in the text data, thus enhancing the performance of ML models through a decrease in data sparsity (Brueckle et al., 2023).

C. Term Representation

This study utilized TF and TF-IDF for feature extraction. The TF is utilized for calculates word frequency within a document (Azam and Yao, 2012) while the TF-IDF is used for word frequency between all documents in a corpus and it is utilized in text data analysis for ML approaches like sentiment analysis and text classification (Martin et al., 2022).

D. Pointwise Mutual Information (PMI)

PMI is a metric utilized to measure the connection between two events within a dataset. It is a common technique in the field of NLP (Ahanin and Ismail, 2022).

This work utilized PMI of significant importance in various domains, but not limited to word sense clarification, information retrieval, and text mining. It enables the detection of significant word relations and can be utilized to create semantic models and extract valuable samples from large text datasets.

The PMI equation is as follows:

$$\mathrm{PMI}\,(x,\ y) = \frac{\log_2 P(x,\ y)}{(P(x) * P(y))} \tag{1}$$

In this equation, P(x, y) is the likelihood of the joint occurrence of events x and y, while P(x) and P(y) are the probabilities of events x and y explaining autonomously.

PMI creates a positive value when the cooccurrence of events x and y is higher than expected, indicating a positive association. A value of 0 signifies independence between the events, while negative values indicate a lower -occurrence than expected, implying a negative association.

E. Proposed Ensemble model

In this study, proposed an ensemble model that combines RF, DT, SVM, and AdaBoost for the detection of ADR is a effective technique for increasing accuracy classifier. Fig. 2 shows the proposed ensemble model proposed. The following describes each of the models used:

DT

DT algorithms are flexible methods for constructing tree-like models using various features and their corresponding thresholds by iterative data partitioning. This algorithm provides a simple and understandable framework for capturing complex patterns within the data, incorporating both categorical and numerical attributes. It is sensitive to overfitting, which causes the use of regularization methods such as reducing to improve this potential point (Alheeti et al., 2023; Charbuty and Abdulazeez, 2021).

RF

RF is a prevalent ensemble ML method appropriate for both classification and regression assignments. Operational on the foundation of DT values, it combines many DTs, each honed on a definite subset of training data, to formulate predictions. Its resilience and adaptability, RF bests in helping high-dimensional dataset brimming with numerous features, rendering it especially accurate in comparison with conventional classification methods (Sheykhmousa et al., 2020; Alsumaidaie et al., 2023).

SVM

SVM is a focus primarily on binary classification, but its means extend to multiclass classification. The core objective of SVM is to decide the best decision limits to maximize the difference between different classes. It is a competent program for dealing with linear separated data and can deal with nonlinear separated data using kernel tricks. SVM is highly regarded for its generalization characteristics and shows effectiveness in high-dimensional characteristic spaces. However, when faced with overlapping or indistinct classes, challenges arise, underlining the importance of the selection of careful kernels and hyperparameters (Cervantes et al., 2020; Alsumaidaie et al., 2023; Bassel et al., 2022).

AdaBoost (adaptive boost)

AdaBoost operates as a set technology that iteratively trains a weak classifier sequence (typically decision boards, characterized by a small structure characterized by only one split) on different subgroups of the training dataset. Each of these weak classifications was designed to emphasize cases that had previously been misclassified or showed a higher error rate. During the training process, AdaBoost assigns more weight to these challenging cases, allowing the later weaker classifiers to focus more on them. The final prediction made by the AdaBoost model is a weighted fusion of the predictions generated by individual weak classifications (Pham et al., 2021).

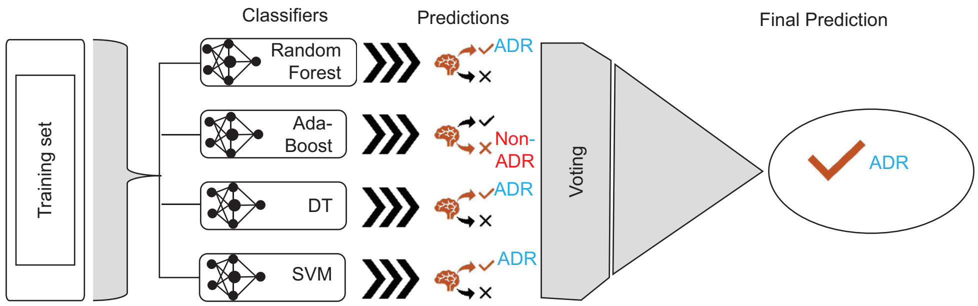

Fig. 2. Proposed ensemble model.

F. Voting

Voting contains each individual classifier in the ensemble making its prediction, and the final prediction is determined by selecting the class that receives the majority of votes. The class with the majority of votes among the individual classifiers is chosen as the final prediction of the ensemble model. The variety of classifiers helps catch different aspects of the data and can lead to more strong predictions.

This proposed utilized combining four algorithms as AdaBoost, SVM, DT, and RF strengths the ensemble model improves classification accuracy of ADR detection and performance, rather than using every algorithm individually. The combined models have improved predictive performance by effectively qualifying bias and errors through combining multiple model outputs. Each ensemble's basic model shows an important role in final prediction and produces more accurate results. Through the combination of RF, SVM, DT, and AdaBoost, the ensemble model develops its unique experiences to focus the several views of ADR detection challenges.

G. Evaluation

There are several performance metrics, such as precision, recall, and F-measure that usually utilized to evaluate the efficiency of ML models.

Precision is used to determine and correctly detect positive cases out of the total predicted positives, showing a low false positive rate (Mukhlif, Al-Khateeb and Mohammed, 2023).

$$\text{Precision} = \frac{\text{True positives (TP)}}{\text{True positives (TP)} + \text{ False positives (FP))}} \quad (2)$$

Recall utilized to evaluates the model's capability to detect all positive cases out of the total actual positives, highlighting a low false negative rate (Kareem and Alheeti, 2022).

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{ False Negatives(FN)}} \quad (3)$$

F-measure is used for merging precision and recall to show a single metric that balances both measures (Alsumaidaie, et al., 2023).

$$F - \text{measure} = \frac{2 \times (\text{Precision } * \text{Recall})}{\text{Precision} + \text{Recall}} \quad (4)$$

IV. Results and Discussion

This section shows the proposed ensemble models results by combining (RF, SVM, and DT with ADA-boost). The objective of the experimentation was to evaluate the effectiveness of the proposed study with a baseline methodology. The baseline approach applied identical data sourced from a benchmark dataset, precisely the annotated ADR review dataset introduced by (Yates et al., 2013). This dataset had been subsequently enriched by Yousef et al. (2019) through the incorporation of supplementary meaningful attributes.

The first baseline proposed LSA with ML algorithms (Nafea, Omar and AL-Ani, 2021), while the second baseline utilized LSA with ANN (Nafea, Omar and Al-qfail, 2023). Both baselines and current study utilized TF or TF-IDF for feature extraction. The testing and training were conducted with the same distribution as the baseline, with 30% of the data allocated for testing and 70% for training.

The first baseline utilized different ML classifiers such as SVM, NB, and LR, while the second baseline used ANN for classifiers. In this study, proposed ensemble models combining RF, SVM, and DT with ADA-boost were applied.

The outcomes of the proposed ensemble models using RF, SVM, and DT with ADA-boost, along with the baseline results, are presented in Table 1. The classification results based on F-measure are shown in Fig. 3, a comparison between the ADR proposed work using ensemble models and baselines research.

The results show the method established, utilizing RF, SVM, DT, and ADA-boost, showed enhancement in F-measure performance when compared with the baseline research. Specifically, when using TF, the proposed ensemble models achieved an enhanced F-measure result of 89%, surpassing the baseline LR result of 82%, while the second baseline using ANN achieved 85%. Similarly, when using TF-IDF, the proposed ensemble models achieved an enhanced F-measure result of 86%, outstanding the baseline LR result of 80%, while the second baseline using ANN achieved 83%.

These results indicate the promising potential of ensemble models, specifically ensemble models, for extracting ADRs. Table III, a comparison between baseline and proposed results.

In addition to the conventional baseline approach, it is essential to explore state-of-the-art methods that apply machine learning or deep learning methods as shown in

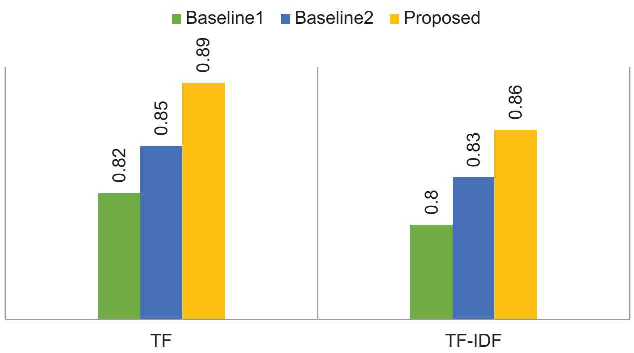

Fig. 3. Comparative of outcomes between the proposed approach and baseline methods.

TABLE III
Comparative of Outcomes between the Proposed Approach and Baseline Methods

Comparison	TF			TF-IDF		
	Precision	Recall	F-measure	Precision	Recall	F-measure
Baseline1	0.83	0.82	0.82	0.82	0.81	0.80
Baseline2	0.85	0.85	0.85	0.83	0.83	0.83
Proposed	0.89	0.89	0.89	0.87	0.86	0.86

TABLE IV
A Comparison of Results between Related Works

Author's	Proposed used	F-measure (%)
Yousef et al. (2019)	SVM, NB and LR with trigger terms w	69
Wang et al. (2019)	DNN	84.4
Zhang et al. (2020)	CNN	67
Zhang et al. (2021)	Bi-LSTM	68
Shen et al. (2021)	GAR	74
Nafea, Omar and AL-Ani (2021)	SVM, NB, and LR with LSA	82
Nafea, Omar and Al-qfail (2023)	ANN with LSA	85
Proposed Method	Ensemble Model (RF+SVM+DT+ADA-boost)	89

DNN: Deep neural network, RF: Random forest, SVM: Support vector machines, DT: Decision trees, NLP: Natural language processing

Table IV. Wang et al. (2019) utilized DNN method and achieved an impressive f-measure of 84.4%. Zhang et al. (2020) developed adversarial transfer learning and applied the private CNN method to identify ADRs, achieving an f-measure of 67%. Zhang et al. (2021) utilized adversarial transfer learning to extract ADRs by using Bi-LSTM, resulting in an f-measure of 68%. In this study by Shen et al. (2021) used a graph adversary representation framework (GAR) that included word embedding. Their approach yielded an f-measure of 74% for the extraction of ADRs. It is essential to acknowledge that making direct comparisons using the provided technique may not be practicable owing to disparities in the dataset used. The efficacy of these deep learning approaches is contingent upon a multitude of parameters, with the size of the dataset being a notable determinant of their success. However, the ensemble model that has been suggested continues to be competitive, especially in situations where small-scale data is being used.

V. Conclusion

The findings of this research using ensemble models using term frequency (TF) achieved an f-measure of 89%, better than the baseline's f-measures of 82% and 85%. This finding shows that when utilizing ensemble models to improve the detection of ADR. Thus, the results show that proposed ensemble models are efficacious in extracting ADR. It is important to acknowledge that this research has several limitations due to its exclusive focus on real-time evaluations. Nevertheless, these issues may play a role in the discovery of novel medication side effects, particularly those associated with COVID-19. In future work, there is potential value in integrating advanced word embedding methodologies with deep learning frameworks. This approach may have been possible to enhance the efficacy of ADR detection.

VI. Appendix

Appendixes, if needed, appear before the acknowledgment.

VII. Acknowledgment

None.

References

Ahanin, Z., and Ismail, M.A., 2022. A multi-label emoji classification method using balanced pointwise mutual information-based feature selection. *Computer Speech and Language*, 73, p.101330.

Alheeti, K.M.A., Alzahrani, A., Alamri, M., Kareem, A.K., and Al_Dosary, D., 2023. A comparative study for SDN security based on machine learning. *International Journal of Interactive Mobile Technologies*, 17(11), pp.131-140.

Alsumaidaie, M.S.I., Alheeti, K.M.A., and Al-Aloosy, A.K., 2023. Intelligent Detection System for a Distributed Denial-of-Service (DDoS) Attack Based on Time Series. In: *2023 15th International Conference on Developments in eSystems Engineering (DeSE)*. IEEE, pp.445-450.

Alsumaidaie, M.S.I., Alheeti, K.M.A., and Alaloosy, A.K., 2023. Intelligent detection of distributed denial of service attacks: A supervised machine learning and ensemble approach. *Iraqi Journal for Computer Science and Mathematics*, 4(3), pp.12-24.

Azam, N., and Yao, J., 2012. Comparison of term frequency and document frequency based feature selection metrics in text categorization. *Expert Systems with Applications*, 39(5), pp.4760-4768.

Bassel, A., Abdulkareem, A.B., Alyasseri, Z.A.A., Sani, N.S., and Mohammed, H.J., 2022. Automatic malignant and benign skin cancer classification using a hybrid deep learning approach. *Diagnostics* (*Basel*), 12(10), p.2472.

Brueckle, M.S., Thomas, E.T., Seide, S.E., Pilz, M., Gonzalez-Gonzalez, A.I., Dinh, T.S., Gerlach, F.M., Harder, S., Glasziou, P.P., and Muth, C., 2023. Amitriptyline's anticholinergic adverse drug reactions-a systematic multiple-indication review and meta-analysis. *PLoS One*, 18(4), p.e0284168.

Cervantes, J., Garcia-Lamont, F., Rodríguez-Mazahua, L., and Lopez A., 2020. A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing*, 408, pp.189-215.

Charbuty, B., and Abdulazeez, A., 2021. Classification based on decision tree algorithm for machine learning. *Journal of Applied Science and Technology Trends*, 2(01), pp.20-28.

Ebrahimi, M., Yazdavar, A.H., Salim, N., and Eltyeb, S., 2016. Recognition of side effects as implicit-opinion words in drug reviews. *Online Information Review*, 40(7), pp.1018-1032.

Edwards, I.R., and Aronson, J.K., 2000. Adverse drug reactions: Definitions, diagnosis, and management. *The Lancet*, 356(9237), pp.1255-1259.

Kareem, A.K., and Alheeti, K.M.A., 2022. Multimodal Approach for Fall Detection based on Support Vector Machine. In: *AIP Conference Proceedings.* AIP Publishing.

Kiritchenko, S., Mohammad, S.M., Morin, J., De Bruijn, B., 2017. NRC-Canada at SMM4H Shared Task: Classifying Tweets Mentioning Adverse Drug Reactions and Medication Intake. In: *CEUR Workshop Proceedings*, p.1-11.

Kiritchenko, S., Zhu, X., and Mohammad, S.M., 2014. Sentiment analysis of short informal texts. *Journal of Artificial Intelligence Research*, 50, pp.723-762.

Li, Z., Yang, Z., Luo, L., Xiang, Y., and Lin, H., 2020. Exploiting adversarial transfer learning for adverse drug reaction detection from texts. *Journal of Biomedical Informatics*, 106, p.103431.

Martin, G.L., Jouganous, J., Savidan, R., Bellec, A., Goehrs, C., Benkebil, M., Miremont, G., Micallef, J., Salvo, F., Pariente, A., and Létinier, L., 2022. Validation of artificial intelligence to support the automatic coding of patient adverse drug reaction reports, using nationwide pharmacovigilance data. *Drug Safety*, 45(5), pp.535-548.

McMaster, C., Chan, J., Liew, D.F.L., Su, E., Frauman, A.G., Chapman, W.W., and Pires, D.E.V., 2023. Developing a deep learning natural language processing algorithm for automated reporting of adverse drug reactions. *Journal of Biomedical Informatics*, 137, p.104265.

Mukhlif, A.A., Al-Khateeb, B., and Mohammed, M., 2023. Classification of breast cancer images using new transfer learning techniques. *Iraqi Journal for Computer Science and Mathematics*, 4(1), pp.167-180.

Nafea, A.A., Omar, N., and AL-Ani, M.M., 2021. Adverse drug reaction detection using latent semantic analysis. *Journal of Computer Science*, 17(10), pp.960-970.

Nafea, A.A., Omar, N. and Al-Qfail, Z.M., 2024. Artificial neural network and latent semantic analysis for adverse drug reaction detection. *Baghdad Science Journal*, 21(1), pp.0226-0233.

Oyebode, O., and Orji, R., 2023. Identifying adverse drug reactions from patient reviews on social media using natural language processing. *Health Informatics Journal,* 29(1).

Pham, B.T., Nguyen, M.D., Nguyen-Thoi, T., Ho, L.S., Koopialipoor, M., Quoc, N.K., Armaghani, D.J., and Van Le H., 2021. A novel approach for classification of soils based on laboratory tests using adaboost, tree and ANN modeling. *Transportation Geotechnics*, 27, p.100508.

Shen, C., Li, Z., Chu, Y., and Zhao Z., 2021. GAR: Graph adversarial representation for adverse drug event detection on Twitter. *Applied Soft Computing*, 106, p.107324.

Sheykhmousa, M., Mahdianpari, M., Mohammadimanesh, F., Ghamisi, P., and Hom, S., 2020. Support vector machine versus random forest for remote sensing image classification: A meta-analysis and systematic review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, pp.6308-6325.

Sørup, F.K.H., Eriksson, R., Westergaard, D., Hallas, J., Brunak, S., and Ejdrup Andersen, S., 2020. Sex differences in text-mined possible adverse drug events associated with drugs for psychosis. *Journal of Psychopharmacology*, 34(5), pp.532-539.

Wang, C.S., Lin, P.J., Cheng, C.L., Tai, S.H., Kao Yang, Y.H., and Chiang, J.H., 2019. Detecting potential adverse drug reactions using a deep neural network model. *Journal of Medical Internet Research*, 21(2), p.e11016.

Yadesa, T.M., Kitutu, F.E., Deyno, S., Ogwang, P.E., Tamukong, R., and Alele, P.E., 2021. Prevalence, characteristics and predicting risk factors of adverse drug reactions among hospitalized older adults: A systematic review and meta-analysis. *SAGE Open Medicine*, 9.

Yates, A., and Goharian, N., 2013. ADRTrace: Detecting expected and unexpected adverse drug reactions from user reviews on social media sites. In: *European Conference on Information Retrieval*. Springer, Berlin, Heidelberg, pp.816-819.

Yousef, R.N., Tiun, S., and Omar, N., 2019. Extended trigger terms for extracting adverse drug reactions in social media texts. *Journal of Computer Science*, 15(6), pp.873-879.

Yousef, R.N.M., Tiun S., Omar N., and Alshari, E.A., 2020. Lexicon replacement method using word embedding technique for extracting adverse drug reaction. *International Journal of Technology Management and Information System*, 2(1), pp.113-122.

Zhang, T., Lin, H., Ren.,Y., Yang, Z., Wang, J., Duan, X., and Xu, B., 2021. Identifying adverse drug reaction entities from social media with adversarial transfer learning model. *Neurocomputing*, 453, pp.254-262.

Zhang, T., Lin, H., Xu, B., Ren, Y., Yang, Z., Wang, J., and Duan, X., 2020. Gated iterative capsule network for adverse drug reaction detection from social media. In: *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, pp.387-390.