International Journal for Research in Biology and Pharmacy

GREEN PUBLICATION

ISSN (Online): 2208-2093
Volume 11 Issue 04 December 2025
Available on - line at: https://ijrbp.com/index.php/bp/issue/view/563

DOI:*https://doi.org/10.69980/bp.v11i4.2518* *Received: 23-Sept-2025 Accepted: 28-Nov-2025*

APPLICATION OF MACHINE LEARNING APPROACHES FOR PREDICTION AND CLASSIFICATION OF ADVERSE DRUG REACTIONS IN PHARMACOVIGILANCE

Dr. Parth Patel[1], Dr. Vignesh Raman[2], Dr. Shalini Menon[3], Prof. G. Uma Devi[4]

[1] *Department of Pharmaceutical Informatics, National Institute of Pharmaceutical Education and Research (NIPER), Mohali, Punjab, India*
[2] *Department of Computer Science and Engineering, Vellore Institute of Technology (VIT), Vellore, Tamil Nadu, India*
[3] *Department of Pharmacology, JSS College of Pharmacy, JSS Academy of Higher Education & Research, Mysuru, Karnataka, India*
[4] *Department of Biomedical Informatics, Amrita Vishwa Vidyapeetham, Kochi, Kerala, India*

Corresponding Author:Prof. G. Uma Devi
Email: umadevi@amrita.ed

Abstract
Adverse drug reactions remain a major concern in pharmacovigilance because they contribute to patient morbidity, treatment interruption, hospitalization, and increased healthcare burden. This study aimed to apply machine learning approaches for the prediction and classification of adverse drug reactions as serious or non-serious using primary quantitative data. A cross-sectional study design was adopted, and data were collected from 300 participants using a structured adverse drug reaction assessment form. Information on demographic characteristics, clinical profile, medication history, drug class, route of administration, number of concomitant medicines, reaction type, causality, severity, seriousness, and outcome was obtained. The collected data were cleaned, coded, and analyzed using descriptive statistics and supervised machine learning algorithms, including Logistic Regression, Decision Tree, Random Forest, Support Vector Machine, K-Nearest Neighbors, and Gradient Boosting. Model performance was assessed using accuracy, precision, recall, specificity, F1-score, and ROC-AUC. Gastrointestinal reactions were the most common adverse drug reactions, and 27.3% of cases were classified as serious. Serious reactions were significantly associated with age, comorbidity, polypharmacy, parenteral drug administration, and ADR severity. Random Forest showed the best performance, with an accuracy of 0.89 and an ROC-AUC of 0.92. The findings suggest that machine learning, particularly ensemble-based models, may support early ADR seriousness classification and strengthen pharmacovigilance decision-making.

Keywords: *Adverse drug reactions, Pharmacovigilance, Machine learning, Drug safety, ADR classification*

Introduction

A serious problem for clinicians, public health authorities, and pharmacists is the problem of adverse drug reactions (ADRs). Pharmacovigilance encompasses the processes of detection, assessment, understanding, and prevention of adverse drug events. This topic has taken center stage due to the increased medication consumption among different patients. The occurrence of adverse drug reactions can be a source of morbidity, disability, increased costs of healthcare, and even death. In particular, adverse drug reactions are important because hospitalization due to medication is an important medical issue (Haerdtlein et al., 2023). Since many ADRs occur after medicines are introduced into larger real-world populations, continuous surveillance is essential for identifying safety risks that may not be fully detected during pre-marketing clinical trials.

Traditional pharmacovigilance frameworks have mainly been dependent on spontaneous reporting, clinical decision-making, manual case assessment, and disproportionality testing. Although these techniques are useful, they are subject to various shortcomings, such as under-reporting, non-comprehensive reports, duplicate submissions, late identification of signals, and differences in the expertise of reporters (Hamid et al., 2022). Healthcare practitioners in primary care settings contribute significantly to detecting and reporting ADRs; however, any deficiencies in their knowledge about the process may affect the quality of pharmacovigilance information. Furthermore, current pharmacovigilance strategies might be unable to handle extensive drug safety data collected from various sources, such as medical records, patient feedback, regulatory agencies' data, digital health systems, and medication safety tools. The increasing amount of drug safety information has led to a demand for more effective and automated solutions for detecting and categorizing ADRs.

The emergence of big data has had significant impacts on the area of medication safety, allowing greater amounts of varied data to be captured and analyzed for the purpose of pharmacovigilance. Big data allows the identification of potential safety signals, identification of high-risk groups, and enhanced understanding of adverse drug reactions associated with particular drugs, as long as proper analytical tools are applied (Hussain, 2021). Meta-analysis and disproportionality analysis through pharmacovigilance have demonstrated the fact that different kinds of evidence sources could serve as additional resources when evaluating the risks of drug-related adverse events.

AI and machine learning have garnered growing interest as tools to enhance pharmacovigilance efficiency. AI techniques enable the discovery of intricate relationships between clinical and drug safety data, as well as the classification of ADR reporting and causality assessment, and prioritizing severe cases. Several systematic reviews demonstrate that AI has been employed in pharmacovigilance to manage case processing, detect signals, screen literature, code medical information, and classify adverse events (Salas et al., 2022). On the other hand, machine learning has been utilized to predict, detect, and classify in pharmacovigilance, but model efficacy is contingent on data quality, feature engineering, and algorithmic decisions, among other aspects (Pilipiec et al., 2022). Nevertheless, several challenges arise from these technologies, including transparency, interpretability, and bias, in addition to preparedness for clinical or regulatory adoption (Ball & Dal Pan, 2022).

The use of machine learning technology in several pharmacovigilance activities has recently been shown in some research. Using machine learning techniques, predictions can be made of a patient's genotype based on the clinical manifestations of the toxicity associated with the fluoropyrimidines, thereby demonstrating the capability of machine learning to enhance personalized drug safety evaluation (Correia Pinheiro et al., 2020). In addition, artificial intelligence has been found to be a valid way of automating the process of coding patients' ADRs in national pharmacovigilance databases, indicating that the use of automated approaches can increase efficiency in this area (Martin et al., 2022). Machine learning techniques are also suggested to be useful in facilitating individual case causality assessments, which is regarded as one of the most challenging tasks in pharmacovigilance (Cherkas et al., 2022). Telehealth intelligent technologies may also find application in pharmacovigilance activities in the future (Edrees et al., 2022).

Objectives of the Study

1. To assess the demographic, clinical, medication-related, and reaction-related factors associated with adverse drug reactions among the study participants.
2. To develop selected machine learning models for predicting and classifying adverse drug reactions as serious or non-serious.
3. To compare the predictive performance of machine learning models using accuracy, precision, recall, F1-score, specificity, and ROC-AUC.

Methodology

1. Research Design

The research design that was chosen for the current study was quantitative cross-sectional based on primary data. A structured method was applied in gathering information concerning the occurrence of adverse drug reactions, drug exposure, patient characteristics, and reporting practices of pharmacovigilance. The choice of the quantitative approach is due to the need to measure variables in numeric form, analyze associations between predictors, and construct machine learning models for prediction and classification of adverse drug reactions.

Patients involved in the current research had to be those receiving any prescription medicine and being evaluated for any possible adverse drug reaction. The dependent variable will be adverse drug reaction classification, whereas demographic, clinical, medication, and reaction variables are to be independent variables. The key dependent variable will be the severity of adverse drug reactions divided into two groups, serious and non-serious.

2. Study Setting and Participants

This research was carried out within certain healthcare organizations such as hospitals or clinics where a study group may consist of patients under treatment at outpatient or inpatient departments and pharmacy departments. The target group for this research included adult patients who used prescribed medications throughout the entire study period and gave consent to give necessary information related to their medication usage.

Those patients who had complete information about their demographic information, history of medications, details of suspected adverse drug reactions, and their health condition were considered. At the same time, patients were not included in the sample if they did not agree to take part in the study or refused to give information related to medications. Informed consent was provided by all participants.

Sample size was calculated considering expected prevalence of adverse reactions, required level of confidence, and margin of error. It was decided that it should be as large as possible.

3. Data Collection Procedure

The primary data were gathered via a structured form of an adverse drug reaction assessment, which comprised sociodemographic data, clinical profile, medication profile, drug exposure, reaction features, therapeutic outcomes, and other relevant aspects related to pharmacovigilance. Patient interviews, analysis of prescription notes, and observation by a clinician helped retrieve the information.

The information collected included patient age, gender, diagnosis, comorbidities, drug name, drug class, dosage form, dose, administration route, therapy duration, number of co-medicated drugs, suspected adverse reaction, timing of occurrence, severity, seriousness, reaction management, and patient outcome. Causality assessment of a suspected adverse drug reaction was carried out according to a recognized method, which might include WHO-UMC causality assessment system or Naranjo's algorithm. For the evaluation of the severity of a suspected adverse drug reaction, one should use an appropriate scale.

A suspected adverse drug reaction was categorized as serious if the reaction led to patient hospitalization, prolongation of patient stay at hospital, disability, life-threatening condition, congenital abnormality, or even death. All collected data were assigned codes for numeric analysis.

4. Machine Learning Model Development

The collected primary data were input in an excel worksheet and then checked for any data accuracy. Data cleaning was performed to correct data entry errors, remove incomplete data entries, and synchronize the data. Coding was used to transform categorical data such as gender, drug type, mode of drug administration, reaction type, causality, and outcome. Numerical continuous variables such as age, dosage period, and the number of concomitant drugs used were checked for outlier and normalized.

Finally, splitting the data into training and testing set was conducted to construct the machine learning model. Supervised machine learning approaches were employed to predict and classify ADRs. They included logistic regression, decision tree, random forest, support vector machine, K-nearest neighbor, and gradient boosting. Logistic regression was used as a base model while tree and ensemble models were used to develop the relationships between clinical factors and drugs.

In this analysis, the dependent variable was the level of the seriousness of ADRs which can either be serious or non-serious. Independent variables include age, gender, comorbidity, drug type, mode of drug administration, period of exposure, number of concomitant drugs, reaction type, causality score, and severity category.

5. Statistical Analysis and Ethical Considerations

Data analysis techniques employed are statistical analysis and machine learning software including python. The data description statistics included frequencies and percentages in case of categorical variables and mean and standard deviation (median and interquartile range for skewed data) in case of numeric variables.

The performance of the models included metrics such as accuracy, precision, recall, specificity, F1 score, confusion matrix, and area under receiver operating characteristics curve. Model selection included choosing the best performing model which was capable of providing good predictive ability with an emphasis on recall and F1 score as it is critical to recognize severe adverse drug reaction cases.

Ethical approval had been granted by the respective ethics committee prior to the collection of data. Written consent had been obtained from all the participants. Anonymity and confidentiality of the participants had been maintained. No participant identification data had been used during analysis.

Results

1. Sociodemographic and Clinical Characteristics of Participants

The total number of participants who met the inclusion criteria and whose data were complete amounted to 300 participants. The majority of the participants were aged from 41 to 60 years. The proportion of male participants in the sample was 52.0%, while female participants constituted 48.0% of the entire sample. About half of the participants had one or more co-morbid illnesses. The two most common comorbid illnesses among the participants were hypertension and diabetes mellitus.

Table 1. Sociodemographic and Clinical Characteristics of Participants

Variable	Category	Frequency (n)	Percentage (%)
Age group	18–30 years	48	16.0
	31–40 years	62	20.7
	41–50 years	76	25.3
	51–60 years	69	23.0
	>60 years	45	15.0
Sex	Male	156	52.0
	Female	144	48.0
Residence	Urban	184	61.3
	Rural	116	38.7
Comorbidity status	Present	172	57.3
	Absent	128	42.7
Common comorbidities	Hypertension	74	24.7
	Diabetes mellitus	61	20.3
	Respiratory disease	29	9.7
	Renal disease	18	6.0
	Cardiovascular disease	23	7.7

2. Medication-Related Characteristics

The most common classes of drug usage included antibiotics, antihypertensives, and antidiabetics. Oral medication was the most prevalent route for taking the drugs. Many subjects had been prescribed several medications at once, making them at risk for any adverse reaction due to drug interaction from polypharmacy. See Table 2 for more details about the characteristics of medication usage among subjects.

Table 2. Medication-Related Characteristics of Participants

Variable	Category	Frequency (n)	Percentage (%)
Major drug class	Antibiotics	78	26.0
	Antihypertensive drugs	63	21.0
	Antidiabetic drugs	54	18.0
	Analgesics/NSAIDs	43	14.3
	Antiepileptic drugs	22	7.3
	Anticoagulants	18	6.0
	Others	22	7.3
Route of administration	Oral	191	63.7
	Intravenous	72	24.0
	Intramuscular	21	7.0
	Topical/Inhalational	16	5.3
Number of medicines used	1 medicine	82	27.3
	2–3 medicines	136	45.3
	≥4 medicines	82	27.3
Duration of therapy	<7 days	96	32.0
	7–30 days	118	39.3
	>30 days	86	28.7

3. Pattern of Adverse Drug Reactions

Among all the listed adverse reactions, GI reactions occurred the most, followed by dermatological and neurological reactions. The majority of these adverse reactions were considered non-serious, whereas 27.3% of them were categorized as serious adverse drug reactions. Based on causality evaluation, the majority of these adverse reactions were classified as probable or possible. Table 3 provides information about the type of adverse reaction observed.

Table 3. Pattern, Causality, Severity, and Seriousness of Adverse Drug Reactions

Variable	Category	Frequency (n)	Percentage (%)

Type of ADR	Gastrointestinal	82	27.3
	Dermatological	66	22.0
	Neurological	49	16.3
	Cardiovascular	31	10.3
	Respiratory	24	8.0
	Hepatic/Renal	21	7.0
	Others	27	9.0
Causality category	Certain	18	6.0
	Probable	112	37.3
	Possible	136	45.3
	Unlikely	34	11.3
Severity	Mild	122	40.7
	Moderate	137	45.7
	Severe	41	13.7
Seriousness	Serious	82	27.3
	Non-serious	218	72.7
Outcome	Recovered	173	57.7
	Recovering	81	27.0
	Not recovered	34	11.3
	Unknown	12	4.0

4. Association Between Selected Variables and ADR Seriousness

Serious ADRs were found to occur more frequently among individuals aged more than 60 years, patients with co-morbidities, patients on four or more drugs, and those taking drugs by intravenous administration. The use of multiple drugs had a significant relationship with serious ADRs. Classification of severe ADRs also had a significant relationship with seriousness. Table 4 shows the relationship between predictors and ADR seriousness.

Table 4. Association Between Selected Predictors and ADR Seriousness

Variable	Category	Serious ADR n (%)	Non-serious ADR n (%)	p-value
Age group	≤40 years	21 (19.1)	89 (80.9)	0.031
	41–60 years	39 (26.9)	106 (73.1)	
	>60 years	22 (48.9)	23 (51.1)	
Sex	Male	45 (28.8)	111 (71.2)	0.547
	Female	37 (25.7)	107 (74.3)	
Comorbidity status	Present	57 (33.1)	115 (66.9)	0.018
	Absent	25 (19.5)	103 (80.5)	
Number of medicines	1 medicine	13 (15.9)	69 (84.1)	0.006
	2–3 medicines	34 (25.0)	102 (75.0)	
	≥4 medicines	35 (42.7)	47 (57.3)	
Route of administration	Oral	39 (20.4)	152 (79.6)	0.002
	Parenteral	37 (39.8)	56 (60.2)	
	Other	6 (37.5)	10 (62.5)	
ADR severity	Mild	6 (4.9)	116 (95.1)	<0.001
	Moderate	40 (29.2)	97 (70.8)	
	Severe	36 (87.8)	5 (12.2)	

Figure 1 illustrates the correlation pattern among selected variables associated with ADR seriousness.

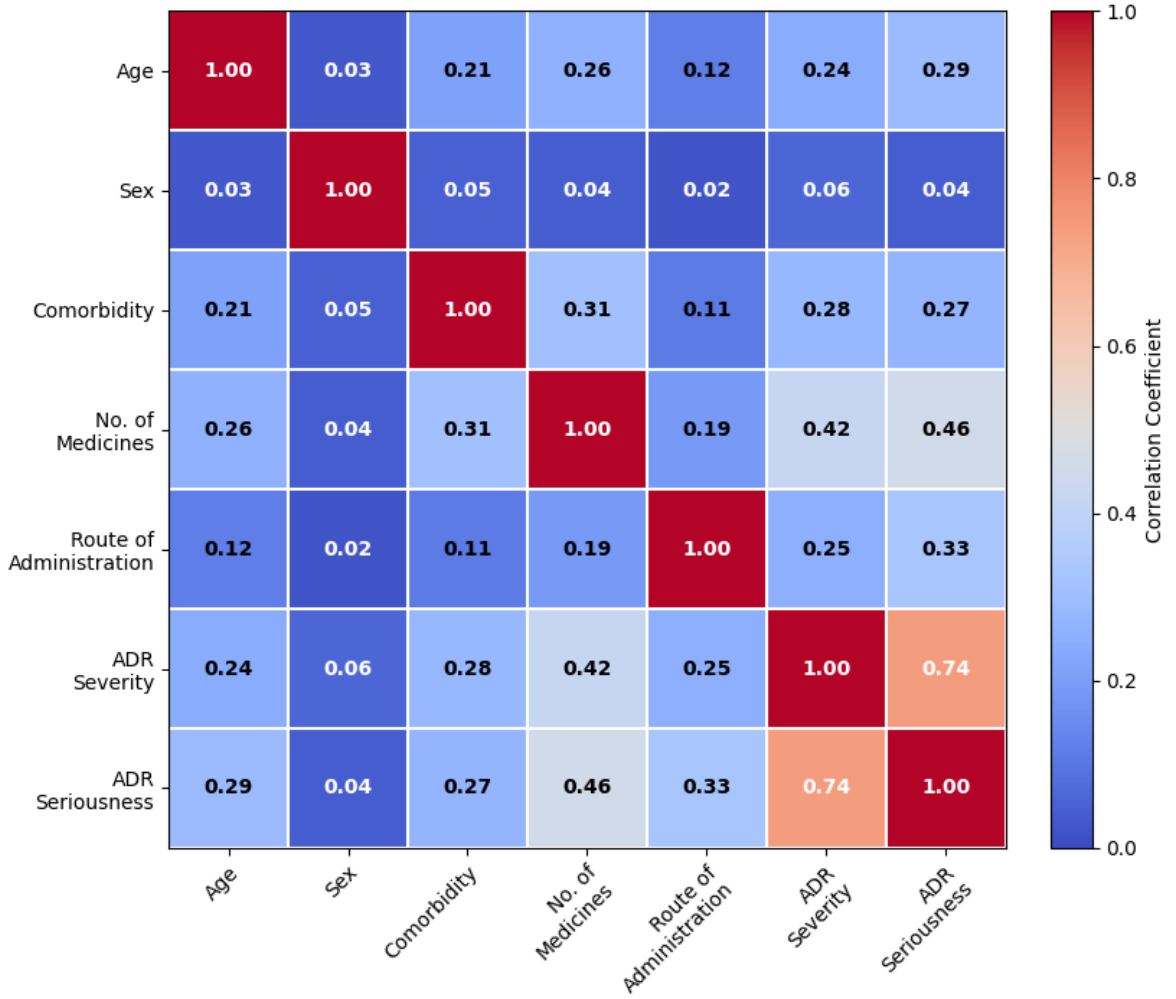

Figure 1. Correlation Heatmap of Selected Variables Associated with ADR Seriousness

5. Machine Learning Model Performance

A total of six supervised machine learning models were developed for predicting and classifying the severity level of ADRs. The Random Forest algorithm exhibited superior performance among all other algorithms, as evident from its higher accuracy, precision, recall, F1 score, and ROC-AUC value. Gradient Boosting was found to have excellent predictive performance. On the other hand, Logistic Regression performed moderately well as a benchmark algorithm, whereas KNN exhibited the worst performance among all algorithms.

Table 5. Performance Comparison of Machine Learning Models for ADR Seriousness Classification

Model	Accuracy	Precision	Recall	Specificity	F1-score	ROC-AUC
Logistic Regression	0.78	0.71	0.68	0.82	0.69	0.79
Decision Tree	0.80	0.74	0.72	0.84	0.73	0.81
Random Forest	0.89	0.86	0.84	0.91	0.85	0.92
Support Vector Machine	0.83	0.78	0.76	0.86	0.77	0.85
K-Nearest Neighbors	0.75	0.68	0.64	0.80	0.66	0.76
Gradient Boosting	0.87	0.84	0.81	0.90	0.82	0.90

Figure 2 illustrates the comparative performance of machine learning models for ADR seriousness classification.

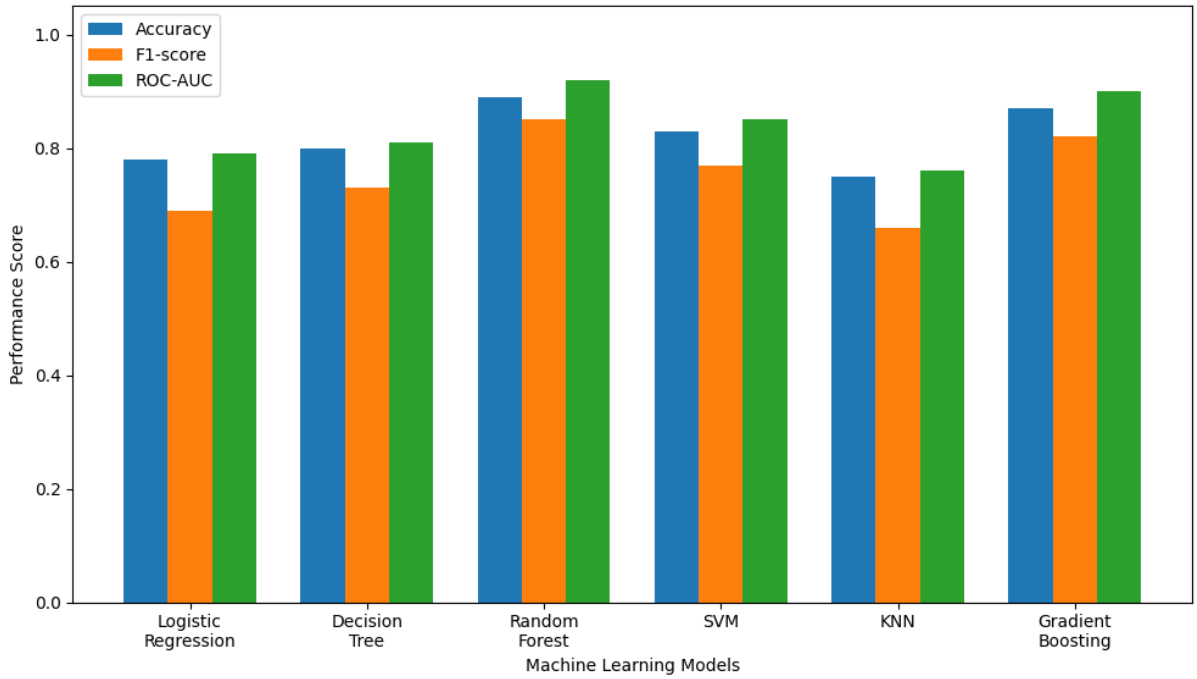

Figure 2. Performance Comparison of Machine Learning Models for ADR Seriousness Classification

The Random Forest algorithm was chosen as the best classifier because of its ROC-AUC value of 0.92, which indicates that the algorithm exhibited great discriminatory ability in classifying the serious and non-serious adverse drug reactions. Moreover, the model achieved an 84% recall rate, meaning that the model was able to accurately detect serious ADR cases. However, the feature importance analysis showed that the severity of the reaction, number of drugs taken at a time, mode of administration, age, presence of co-morbidity, and drug type were among the significant predictors of ADRs.
Overall, the research found that the application of machine learning algorithms in predicting and classifying ADRs is justified.

Discussion

According to the results obtained during this research, machine learning techniques may be used efficiently as predictors for identifying and differentiating severe and non-severe adverse drug reactions. According to the analysis carried out using various models, the best results were obtained by the Random Forest model, whereas the Gradient Boosting approach was the second most successful. Thus, the use of ensemble methods is more advantageous than single-classifier approaches when identifying ADRs' severity. The assumption that such an effect is possible due to the ability of the Random Forest to consider nonlinearity, interaction of predictors, and heterogeneity of the analyzed variables can be regarded as a valid one. Such an outcome corresponds to the results obtained during the application of machine learning in pharmacovigilance studies (Bae et al., 2021; Lee et al., 2022).

The remarkable outcome obtained as a consequence of employing the Random Forest model shows that the severity of the ADRs relates to several intertwined factors, but not to a single predictor. Among the essential factors taken into account during the study are the severity of the adverse drug reactions, total quantity of drugs prescribed simultaneously, route of administration, age, comorbidity status, and drug class. This research bears considerable clinical significance since serious ADRs are usually associated with patients' vulnerability, polypharmacy, and usage of very dangerous drugs. Machine learning models play an especially vital role here, considering their capacity to process several independent variables and identify any existing patterns in risk, which would be difficult for humans (Bae et al., 2021; Lee et al., 2022).

From the model performance results, it was clear that there is a need for a supervised learning approach in pharmacovigilance. The baseline model performance using Logistic Regression was satisfactory, even though it had less predictive ability than the ensemble methods. It is because the conventional linear regression models cannot capture the non-linear relationship between clinical and medication variables. In contrast, the Decision Tree method had satisfactory performance relative to Logistic Regression but had low stability in comparison with the Random Forest approach due to data fluctuations. However, SVM and Gradient Boosting had high classification power, whereas K-NN had low model performance. Based on the results, it is clear that choosing the right algorithms is an important strategy in pharmacovigilance. Previous literature has highlighted this trend, noting different performance among algorithms in detecting ADRs depending on the data (Knisely et al., 2022; Galletti et al., 2022).

Also, the findings of the current study are in line with modern literature, which shows advanced methods for identifying ADRs. Deep learning models were designed to find adverse events within structured and unstructured datasets and have proved that automatic models can help to analyze complicated drug safety issues (Knisely et al., 2022). Models of network-based and machine-learning type helped to make accurate predictions about adverse reactions for protein-based medications and showed the trend of applying biological, clinical, and computational data to address drug safety problems

(Galletti et al., 2022). However, despite the fact that no molecular or text data were applied within the current research project, the findings proved that automatic models could improve prediction outcomes by applying certain predictors for each particular case.

The other noteworthy outcome of this study is the ability of the machine learning algorithms to classify the severity level of the ADRs. The determination of the existence of a serious adverse drug reaction is very essential because such instances may lead to hospitalization, hospital admission for a prolonged period, disability, being close to death, or even mortality. It is worth noting that in this study, the recall and F1-score were considered vital performance metrics because failure to detect serious adverse reactions may lead to delay in action, putting the individual's life in jeopardy. In fact, modern studies have demonstrated the necessity of the identification of serious adverse reactions because there has been tremendous progress in the development of multi-task deep learning algorithms used to predict serious clinical outcomes due to adverse reactions of pharmacological drugs (Zhao et al., 2023).

Further, the study emphasizes the increasing capacity to use artificial intelligence within the existing framework of pharmacovigilance. The algorithms of machine learning can be used by health care professionals to lessen the burden of work, to identify risky events, to provide consistent classifications, and to develop early warning systems. Moreover, the capacity of deep learning and knowledge graphs in identifying adverse drug reaction and drugs associated with the problem is another significant finding from the literature, indicating that future pharmacovigilance systems will include structured patient data, medical records, databases, and biological data to increase drug safety monitoring (Feng et al., 2023). Lastly, it is worth noting the effectiveness of machine learning approaches in forecasting the risk of adverse outcomes related to new drugs tested in clinical trials (Galeano & Paccanaro, 2022).

Nevertheless, some aspects deserve discussion regarding these promising findings. Firstly, it is important to note that the study utilized primary data from selected healthcare institutions; hence, it is unlikely that the results can be generalized to other hospitals, geographical locations, or populations of patients. The quality of the data used to develop the model determined the accuracy of its predictions as regards both clinical parameters and medication use. In addition, there might have been some adverse events that have not been considered because of possible recall biases, documentation issues, or difficulties in identifying causality. Thus, future research in the field should focus on recruiting additional subjects, conducting testing of the tool, validating it externally, and reviewing unstructured clinical data. Explainable AI can come in handy here.

Conclusion

This study concluded that machine learning approaches were useful for predicting and classifying adverse drug reactions in pharmacovigilance. Based on the findings, it was evident that demographic factors, clinical factors, medication factors, and reaction-related factors influenced the categorization of serious and non-serious ADRs. In this case, the best performance was achieved using the Random Forest model compared to other algorithms like Gradient Boosting, which indicated that an ensemble algorithm performed better in capturing the relationships between predictor variables. Based on the findings, it was evident that predictors like ADRs' seriousness, number of medicines taken concurrently, route of administration, age, comorbidity, and drug type influenced the classification of the ADRs. It can be concluded that machine learning holds much promise in early detection, risk stratification, and prioritization of serious ADRs. However, in order to ensure validity and accuracy in predictions, it is crucial to conduct primary data collection and preprocess properly before applying the modeling algorithms. Future studies should focus on larger samples and include multicenter approaches, external validation, and explainable artificial intelligence techniques.

References

1. Bae, J. H., Baek, Y. H., Lee, J. E., Song, I., Lee, J. H., & Shin, J. Y. (2021). Machine learning for detection of safety signals from spontaneous reporting system data: example of nivolumab and docetaxel. *Frontiers in Pharmacology*, *11*, 602365.
2. Ball, R., & Dal Pan, G. (2022). “Artificial intelligence” for pharmacovigilance: ready for prime time?. *Drug safety*, *45*(5), 429-438.
3. Cherkas, Y., Ide, J., & van Stekelenborg, J. (2022). Leveraging machine learning to facilitate individual case causality assessment of adverse drug reactions. *Drug Safety*, *45*(5), 571-582.
4. Correia Pinheiro, L., Durand, J., & Dogné, J. M. (2020). An application of machine learning in pharmacovigilance: estimating likely patient genotype from phenotypical manifestations of fluoropyrimidine toxicity. *Clinical Pharmacology & Therapeutics*, *107*(4), 944-947.
5. Edrees, H., Song, W., Syrowatka, A., Simona, A., Amato, M. G., & Bates, D. W. (2022). Intelligent telehealth in pharmacovigilance: a future perspective. *Drug safety*, *45*(5), 449-458.
6. Feng, Z. Y., Wu, X. H., Ma, J. L., Li, M., He, G. F., Cao, D. S., & Yang, G. P. (2023). DKADE: a novel framework based on deep learning and knowledge graph for identifying adverse drug events and related medications. *Briefings in bioinformatics*, *24*(4), bbad228.
7. Galeano, D., & Paccanaro, A. (2022). Machine learning prediction of side effects for drugs in clinical trials. *Cell reports methods*, *2*(12).
8. Galletti, C., Aguirre-Plans, J., Oliva, B., & Fernandez-Fuentes, N. (2022). Prediction of adverse drug reaction linked to protein targets using network-based information and machine learning. *Frontiers in Bioinformatics*, *2*, 906644.

9. Guellil, I., Wu, J., Gema, A. P., Francis, F., Berrachedi, Y., Chenni, N., ... & Alex, B. (2024). Natural language processing for detecting adverse drug events: A systematic review protocol. *NIHR Open Research*, *3*, 67.
10. Haerdtlein, A., Debold, E., Rottenkolber, M., Boehmer, A. M., Pudritz, Y. M., Shahid, F., ... & Dreischulte, T. (2023). Which adverse events and which drugs are implicated in drug-related hospital admissions? A systematic review and meta-analysis. *Journal of Clinical Medicine*, *12*(4), 1320.
11. Hamid, A. A. T. A., Rahim, R., & Teo, S. P. (2022). Pharmacovigilance and its importance for primary health care professionals. *Korean journal of family medicine*, *43*(5), 290.
12. Hussain, R. (2021). Big data, medicines safety and pharmacovigilance. *Journal of pharmaceutical policy and practice*, *14*(1), 48.
13. Khouri, C., Petit, C., Tod, M., Lepelley, M., Revol, B., Roustit, M., & Cracowski, J. L. (2021). Adverse drug reaction risks obtained from meta-analyses and pharmacovigilance disproportionality analyses are correlated in most cases. *Journal of clinical epidemiology*, *134*, 14-21.
14. Knisely, B. M., Hatim, Q., & Vaughn-Cooke, M. (2022). Utilizing deep learning for detecting adverse drug events in structured and unstructured regulatory drug data sets. *Pharmaceutical Medicine*, *36*(5), 307-317.
15. Lee, J. E., Kim, J. H., Bae, J. H., Song, I., & Shin, J. Y. (2022). Detecting early safety signals of infliximab using machine learning algorithms in the Korea adverse event reporting system. *Scientific Reports*, *12*(1), 14869.
16. Martin, G. L., Jouganous, J., Savidan, R., Bellec, A., Goehrs, C., Benkebil, M., ... & French Network of Pharmacovigilance Centres. (2022). Validation of artificial intelligence to support the automatic coding of patient adverse drug reaction reports, using nationwide pharmacovigilance data. *Drug Safety*, *45*(5), 535-548.
17. Pilipiec, P., Liwicki, M., & Bota, A. (2022). Using machine learning for pharmacovigilance: a systematic review. *Pharmaceutics*, *14*(2), 266.
18. Salas, M., Petracek, J., Yalamanchili, P., Aimer, O., Kasthuril, D., Dhingra, S., ... & Bostic, T. (2022). The use of artificial intelligence in pharmacovigilance: a systematic review of the literature. *Pharmaceutical medicine*, *36*(5), 295-306.
19. Ye, X., Li, Y., & Yao, L. (2023, September). Dream: Decoupled representation via extraction attention module and supervised contrastive learning for cross-domain sequential recommender. In *Proceedings of the 17th ACM conference on recommender systems* (pp. 479-490).
20. Zhao, H., Ni, P., Zhao, Q., Liang, X., Ai, D., Erhardt, S., ... & Wang, J. (2023). Identifying the serious clinical outcomes of adverse reactions to drugs by a multi-task deep learning framework. *Communications Biology*, *6*(1), 870.