JOURNAL OF PHARMA INSIGHTS AND RESEARCH

REVIEW ARTICLE

A Systemic review of machine learning approaches for adverse drug reaction detection: Novel perspective and challenges

Suberna Basnet [1]*, Ali Nihal [1], Sijina KS [1], Naga Kireeti Seru [1], Amit Kumar [2]
[1] *PharmD Intern, Department of Pharmacy Practice, Aditya College of Pharmacy, Surampalem, AP, India*
[2] *Associate Professor, Department of Pharmacy Practice, Aditya College of Pharmacy, Surampalem, AP, India*

Publication history: Received on 28th October; Revised on 25th November; Accepted on 30th November

Article DOI: 10.5281/zenodo.10289842

Abstract: Medication errors significantly impact patient treatment outcomes, necessitating the integration of modern technologies for improved detection and prevention. This review amalgamates findings from multiple studies on medical decision support systems and machine learning to predict and mitigate prescribing errors. A systematic examination of 30 articles published between 2015 and 2023 reveals the utilization of various methodologies, including outlier detection testing, interruptive prescribing alerts, and probabilistic, machine learning-based clinical decision support systems. The review underscores the imminent need for sophisticated techniques to address the limitations of traditional Adverse Drug Reaction (ADR) detection methods. Notably, the incorporation and refinement of machine learning approaches emerge as promising strategies. The examination of these studies highlights the potential of machine learning to revolutionize patient safety and healthcare quality by enhancing efficiency and accuracy. In conclusion, this review emphasizes that machine learning represents a groundbreaking approach in detecting and preventing medication errors. The integration of advanced methods, coupled with a robust reporting system, is crucial for advancing the landscape of ADR discovery. This approach not only facilitates efficient and accurate healthcare delivery but also ensures a patient-centric focus, marking a significant stride towards improved patient safety and healthcare quality.

Keywords: Machine learning; Adverse drug reactions; Clinical decision support systems; Medication errors; Patient safety

1. Introduction

Pharmaceutical formulations consist of chemical compounds designed to elicit therapeutic responses against pathological conditions; however, instances may arise wherein these chemical agents manifest paradoxical effects, leading to morbidity or fatality.[1] ADR management costs can reach $250 US dollars for a person in India, which is unnecessarily costly for a developing nation where the average income of lower class people is less than $100 US dollars. India ADR reporting rate is less than 1%, whereas the global figure is 5%. ADR occurrence is statistically analyzed with the aid of methods such as the spontaneous reporting system, which is used for signal detection.[2] However, there are a number of limitations and issues with SPS, including underreporting and bias in the identification of pharmacological side effects. The diagnosis of adverse events is frequently not made clear in professional medical reports of adverse occurrences. In actuality, even though the majority of ADRs are on the list of differential diagnoses that physicians can use, it might be challenging to diagnose them.[3] Generally speaking, if the cause is not obvious, it is not reported as an ADR. As a result, numerous investigations taking into account this SRS constraint are currently underway. Electronic medical record (EMR) data are another source of ADR study data, and they are crucial for verifying clinical evidence. [4]

They furnish detailed statistical delineations concerning the experiences of afflicted patients undergoing medical interventions, encompassing prognostic assessments and the duration of prescribed medications, commencing from the initiation to the cessation dates of the prescribed drug regimen [5] Timely detection of ADEs could allow screening and the diagnosis of adverse events regularly which isn't made clean in professional clinical reports of adverse event. Additionally, it is tough to diagnose ADR in spite of the fact that the ones patients are below monitoring. [6]

The process involves integrating data from several sources obtain from the presents significant problems for traditional reporting systems such as the prescription event monitoring, chart review, spontaneous reporting system, and medication report. Here the problem with those systems is prone to discrepancies. Since they depend on continuous reporting and prescription patterns. Those shortcomings in classic ADR detection techniques include underreporting, costly manual procedures, and difficulties integrating

* Corresponding author: Suberna Basnet

data. These shortcomings have encouraged the researches to develop the advanced methods such as machine learning to improve the effectiveness of ADR detection. [7].

2. Methodology

Our systematic review, following PRISMA guidelines, targeted clinical research articles from January 2015 to March 2023, focusing on adverse drug reactions (ADRs) and related methods. The search spanned EMBASE and PubMed, including terms like adverse events, drug safety, statistical analysis, and machine learning. The inclusion criteria encompassed original articles with relevant terms in titles and abstracts. Exclusion criteria eliminated studies outside the stipulated timeframe or those not aligned with the review's objectives. Risk of bias was assessed using the ROBIS tool, evaluating randomization, intervention deviation, missing data, outcome measurement, and selection bias. The analysis methods were categorized into statistical (disproportionate analysis, regression, Log-likelihood ratio test) and machine learning (Bayesian methods, supervised methods, other methods as shown in Figure 1). The bias evaluation process was applied to both statistical and machine learning methodologies. This comprehensive approach aims to provide a rigorous and unbiased synthesis of recent trends in ADR detection methods.[8]

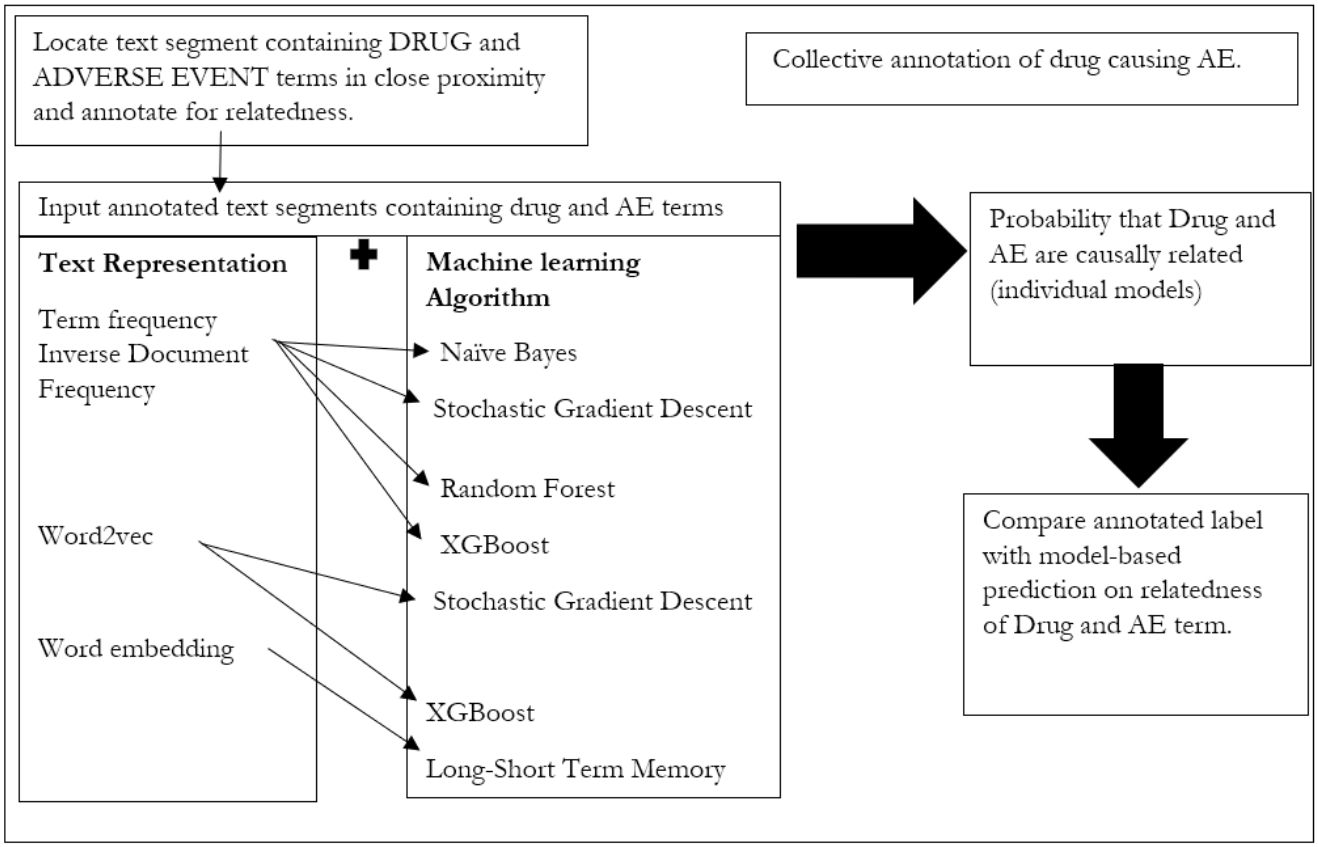

Figure 1. Showing Machine Learning Approaches in AE Detection

3. Results

A systemic analysis was conducted on number of distinct publications categorized as machine learning and statistical approaches respectively. The study presented in those articles is based on data sources, including FAERS, SIDER, EU databases, SRSs, and other national databases. ML algorithms, such as naïve bayes and random forests were apply to these databases in order to find the probability of AE. Drug-drug and drug-disease interaction were analysed to find out medication related responses as AE

3.1 Statistical methods for ADR detection

The statistical methods employed for ADR detection were predominantly based on spontaneous adverse event reports collected through voluntary reporting systems (SRS) and electronic medical record (EMR) data. Disproportionate analysis, log-likelihood ratio tests (LRT), and regression methods, such as survival and logistic regression, were extensively utilized. Notably, >70% of the results

were associated with statistical analyses using SRS and EMR data (Table 1). The FDA's adverse event reporting system (FAERS) and VigiBase by the World Health Organization were among the major sources, with studies assessing various aspects of drug safety and ADR signals. The application of LRT in SRS data was less frequent but notable, with studies focusing on temporal variations and heterogeneity in reporting rates.

Table:1. Statistical Methods of ADR Detection in SRS data

Statistical methods	Source	Outcome
Signals of disproportionate reporting (SDRs) [9]	SRS data	Different algorithms performed differently between databases but Performance was strongly dependent on the thresholds used to define a statistical signal.
Disproportionality[10]	EV database	Continues evaluation of old and new ADR data to evaluate the new/old drug outcome.
Disproportionality[11]	FAERS	Proportion of the signals for "serious" events was significantly higher in the group of non-signals by ROR

3.2 Machine learning methods for ADR detection

Machine learning methods were diverse, categorized into Bayesian methods, supervised methods, and other methods. Studies focused on various data sources, including SRS, EMR, and other databases like DrugBank, PubChem, and specific disease-related datasets.[12]

3.3 Machine Learning Methods for ADR Detection in SRS Data

Supervised methods such as random forest, support vector machine, and recurrent neural networks were employed for ADR detection in SRS data. These studies aimed to enhance signal detection and predict ADRs associated with specific drugs, considering sex differences, vaccine safety, and injection-related ADRs in children.

3.4 Machine Learning Methods for ADR Detection in Other Data Sources

Supervised methods, such as random forest, pairwise kernel SVM classifiers, and XGBoost algorithms, were applied to various data sources beyond SRS (Table 2). These studies aimed to predict unknown drug-ADR pairs, identify ADRs caused by lead compounds, and predict the risk of analgesic side effects.[13]

Table 2. Machine Learning Methods of ADR Detection Using Various Source of Data.

ML methods	Source of data	Objective
Bayesian[14]	FAERS, MedEffecr	To determine the drug safety from multiple data source
Inductive matrix completion [15]	Drugbank, FAERS	To minimize the drug induce functional loss
Blockmetrices [16]	VAERS	To use correct matrix for detect AE of vaccine.
Unsupervised and supervised method [17]	Drugbank, KEGG, NDF-RT	To define the process if detecting DDI.
Feature-derived graph [18]	SIDER & EMR	To improve the knowledge of drug induce ADR
Random foorest [19]	EMR,DrugBank	Improvising data mining method for ADR detection.

4. Discussion

4.1. Machine learning in ADR detection

Machine learning (ML) has become a transformative tool in healthcare, offering innovative solutions to various challenges, including the detection of Adverse Drug Reactions (ADRs). In healthcare, ML leverages algorithms and statistical models can be applied to analyse vast datasets, providing valuable insights for diagnosis, treatment and defiantly ADR detection. The transform towards machine learning in ADR detection helps to result into more efficient and accurate methodologies.[8] Traditional approaches, only depends on manual processes and structured reporting which have huge limitations such as underreporting and time-consuming workflows. Machine learning introduces a paradigm shift by automating the analysis of complex healthcare data, enabling scalable and real-time ADR detection. The integration of ML into healthcare systems improve patient safety through the timely identification of potential adverse events. [20]

Automation significantly reducing the manual effort involved in traditional methods. ML models can autonomously analyse large volumes of electronic health records, clinical notes, and other healthcare data sources, expediting the identification of patterns indicative of potential ADRs. Scalability is another notable advantage as ML models can efficiently process extensive datasets, accommodating the growing volume of healthcare information generated daily. Further ML contributes to improved accuracy by discerning subtle relationships within data that might be challenging for human observers. These models can learn from historical data, adapt to evolving patterns, and enhance predictive capabilities, thereby revolutionizing ADR detection in terms of efficiency and effectiveness. [21]

4.2. Current challenges in ADR detection

ADR detection faces significant challenges with machine learning that why careful consideration should be carried out. One major challenge is data heterogeneity, stemming from the diverse sources and formats of healthcare data. ML models must contend with variations in data structures, terminologies, and quality, necessitating robust pre-processing strategies to ensure the uniformity and reliability of input data. Imbalanced datasets pose another challenge, where the occurrence of ADRs may be relatively rare compared to non-events. This disbalance can lead to biased models that prioritize the majority class, potentially overlooking critical ADR signals. Addressing imbalanced datasets requires specialized techniques such as oversampling minority classes or adjusting class weights during model training. [22]

Interpretability remains a crucial challenge in ML-based ADR detection. While these models can generate accurate predictions, understanding the rationale behind their decisions is often complex. The "black box" nature of some ML algorithms poses challenges in gaining clinicians' trust and acceptance. [23] As healthcare decisions heavily rely on interpretability, efforts to enhance the transparency and explain ability of ML models are paramount. Additionally, incorporating domain knowledge and expert insights into the ML process can help bridge the gap between sophisticated algorithms and the practical requirements of healthcare professionals.

While machine learning brings transformative potential to ADR detection, addressing challenges such as data heterogeneity, imbalanced datasets, and interpretability issues is essential for ensuring the reliability and acceptance of ML-based approaches in healthcare. Efforts to overcome these challenges will contribute to the integration of machine learning into routine clinical practice, ultimately enhancing the safety and quality of patient care. Traditional approaches apply on spontaneous reporting systems (SRS) and manual processes, face inherent challenges such as underreporting and time-consuming workflows. Machine learning (ML) technology act as the foundation of innovation, offering automation, scalability, and improved accuracy.

The prevalence of statistical methods, particularly those utilizing SRS and electronic medical record (EMR) data, for ADR detection. Disproportionate analysis, log-likelihood ratio tests (LRT), and regression methods have been instrumental in exploring drug safety profiles and detecting ADR signals. Machine learning methods, categorized into Bayesian, supervised, and other approaches, exhibit diverse applications, ranging from predicting unknown drug-ADR pairs to identifying ADRs caused by lead compounds.

5. Conclusion

In conclusion, the integration of machine learning into routine clinical practice holds immense promise for enhancing patient safety through the timely identification of potential adverse events. The synthesis of advanced methodologies, combined with a robust understanding of the evolving landscape of ADR detection, offers a pathway toward more efficient, accurate, and patient-centred healthcare. Such scenario may not be possible to today healthcare system of developing country like ours but with the development of innovation and electronic prescription followed by collecting and storing individual data in the vast system network will ultimately help in the development on site specific ML application on the individual patient care process.

Abbreviations

ADR-Adverse Drug Reaction

SPS- Spontaneous Reporting System

EMR- Electronic Medical Record

ML - Machine Learning

PRISMA- Preferred Reporting Items for Systematic Reviews and Meta-Analyses

EMBASE- Excerpta Medica dataBASE

ROBIS tool- Risk of Bias In Systematic review

EU database - European Union database

LRT- Log-Likelihood Ratio Tests

FAERS-FDA Adverse Event Reporting System

SIDER-Side Effect Resource

VAERS- Vaccine Adverse Event Reporting System

References

[1] Lee CY, Chen YP. Machine learning on adverse drug reactions for pharmacovigilance. Drug discovery today. 2019 Jul 1;24(7):1332-43.

[2] Yang S, Kar S. Application of artificial intelligence and machine learning in early detection of adverse drug reactions (ADRs) and drug-induced toxicity. Artificial Intelligence Chemistry. 2023 Aug 10:100011.

[3] Kongkaew C, Noyce PR, Ashcroft DM. Hospital admissions associated with adverse drug reactions: a systematic review of prospective observational studies. Annals of Pharmacotherapy. 2008 Jul;42(7-8):1017-25.

[4] Denck J, Ozkirimli E, Wang K. Machine-learning-based adverse drug event prediction from observational health data: a review. Drug Discovery Today. 2023 Jul 17:103715.

[5] Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, Van Calster B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. Journal of clinical epidemiology. 2019 Jun 1;110:12-22.

[6] Vamathevan J, Clark D, Czodrowski P, Dunham I, Ferran E, Lee G, Li B, Madabhushi A, Shah P, Spitzer M, Zhao S. Applications of machine learning in drug discovery and development. Nature reviews Drug discovery. 2019 Jun;18(6):463-77.

[7] Huang K, Fu T, Gao W, Zhao Y, Roohani Y, Leskovec J, Coley CW, Xiao C, Sun J, Zitnik M. Therapeutics data commons: Machine learning datasets and tasks for drug discovery and development. arXiv preprint arXiv:2102.09548. 2021 Feb 18.

[8] Kavakiotis I, Tsave O, Salifoglou A, Maglaveras N, Vlahavas I, Chouvarda I. Machine learning and data mining methods in diabetes research. Computational and structural biotechnology journal. 2017 Jan 1;15:104-16.

[9] Pang G, Shen C, Cao L, Hengel AV. Deep learning for anomaly detection: A review. ACM computing surveys (CSUR). 2021 Mar 5;54(2):1-38.

[10] Chen S, Li T, Yang L, Zhai F, Jiang X, Xiang R, Ling G. Artificial intelligence-driven prediction of multiple drug interactions. Briefings in Bioinformatics. 2022 Nov;23(6):bbac427.

[11] Ventola CL. Big data and pharmacovigilance: data mining for adverse drug events and interactions. Pharmacy and therapeutics. 2018 Jun;43(6):340.

[12] Kumar SA, Ananda Kumar TD, Beeraka NM, Pujar GV, Singh M, Narayana Akshatha HS, Bhagyalalitha M. Machine learning and deep learning in data-driven decision making of drug discovery and challenges in high-quality data acquisition in the pharmaceutical industry. Future Medicinal Chemistry. 2022 Feb;14(4):245-70.

[13] Ibrahim MS, Dong W, Yang Q. Machine learning driven smart electric power systems: Current trends and new perspectives. Applied Energy. 2020 Aug 15;272:115237.

[14] Sokolova M, Lapalme G. A systematic analysis of performance measures for classification tasks. Information processing & management. 2009 Jul 1;45(4):427-37.

[15] Mann K, Gordon J, MacLeod A. Reflection and reflective practice in health professions education: a systematic review. Advances in health sciences education. 2009 Oct;14:595-621.

[16] Mann K, Gordon J, MacLeod A. Reflection and reflective practice in health professions education: a systematic review. Advances in health sciences education. 2009 Oct;14:595-621.

[17] Islam MS, Hasan MM, Wang X, Germack HD, Noor-E-Alam M. A systematic review on healthcare analytics: application and theoretical perspective of data mining. InHealthcare 2018 May 23 (Vol. 6, No. 2, p. 54). MDPI.

[18] Farrar CR, Worden K. Structural health monitoring: a machine learning perspective. John Wiley & Sons; 2012 Nov 19.

[19] Arjmand B, Hamidpour SK, Tayanloo-Beik A, Goodarzi P, Aghayan HR, Adibi H, Larijani B. Machine learning: a new prospect in multi-omics data analysis of cancer. Frontiers in Genetics. 2022 Jan 27;13:824451.

[20] Mullainathan S, Spiess J. Machine learning: an applied econometric approach. Journal of Economic Perspectives. 2017 May 1;31(2):87-106.

[21] Kendale S, Kulkarni P, Rosenberg AD, Wang J. Supervised machine-learning predictive analytics for prediction of postinduction hypotension. Anesthesiology. 2018 Oct 1;129(4):675-88.

[22] Zhao ZQ, Zheng P, Xu ST, Wu X. Object detection with deep learning: A review. IEEE transactions on neural networks and learning systems. 2019 Jan 27;30(11):3212-32.

[23] Ballantine J, Stuber J, Everitt J. The sociology of education: A systematic analysis. Routledge; 2021 Jul 29.

[24] Lundervold AS, Lundervold A. An overview of deep learning in medical imaging focusing on MRI. Zeitschrift für Medizinische Physik. 2019 May 1;29(2):102-27.

Author's short biography

Suberna Basnet:

Suberna Basnet is an aspiring Pharm-D student with the passion of optimising patient healthcare through comprehensive pharmaceutical care. Currently pursuing Pharm-D with a focus on drug therapy optimization, patient counselling. Eager to apply theoretical knowledge in real world clinical settings. Demonstrates strong work ethics, adaptability and a commitment to ongoing professional development in the field of clinical pharmacy

Seru Naga Kireeti:

Naga Kireeti is a committed PharmD student involved in providing clinical pharmacy services. Enthusiastically exploring the nuances of clinical pharmacy practice, Naga also fosters a strong interest in research pursuits. With a fervour for educating patients about medications and a sharp focus on drug dosing and interactions and a sharp focus on drug dosing and interactions. Naga aims to make a significant contribution to pharmacy field

Ali Nihal:

Ali Nihal, an ardent pharm-D Student, actively participates in delivering clinical pharmacy services. Exploring the intricate domains of clinical pharmacy practice. He nurtures a profound interest in research pursuits. With a passion for educating patients on medications and a dedicated focus on drug dosing and interactions, he aspires to bring valuable contribution to the field of pharmacy

Sijina KS:

Sijina KS, a 23-year-old Pharm D intern at Aditya College of Pharmacy. With a strong foundation in patient counselling, medical history collection, and computer technology, she excel in project management, team leadership, active listening, empathy, self-management, and possess excellent communication skills. Her ability to handle diverse patients in critical situations and effective presentation skills stand out. She have actively participated in various National seminars and published a case report in 2022

Amit Kumar:

Amit Kumar is an accomplished professional in the field of pharmacy, holding a B.Pharmacy, M. Pharmacy, and had submitted his Ph.D. Currently serving as the Associate Professor and Head of the Pharmacy Practice Department at the NAAC A accredited Aditya College of Pharmacy in Surampalem. He has demonstrated his commitment to advancing pharmaceutical knowledge through his extensive publication record, with 33 articles published in various reputed Indian and international journals. His research contributions span a range of topics within the pharmaceutical domain, showcasing his expertise and dedication to the field.