International Journal of Engineering Technology and Management Science
Website: ijetms.in Issue: 2 Volume No.9 March - April – 2025
DOI:10.46647/ijetms.2025.v09i02.096 ISSN: 2581-4621

Explainable Artificial Intelligence for Patient Safety A Review of Application in Pharmacovigilance

Thalari Kalyan Babu[1], CH. Sri Lakshmi Prasanna[2]
[1]*MCA Student, Dr.K.V.Subba Reddy Institute of Technology, Kurnool, Andhra Pradesh, India*
[2]*Assistant Professor, Dr.K.V.Subba Reddy Institute of Technology, Kurnool, Andhra Pradesh, India*

Abstract
Explainable AI (XAI) is a methodology that complements the black box of artificial intelligence, and its necessity has recently been highlighted in various fields. The purpose of this research is to identify studies in the field of pharmacovigilance using XAI. Though there have been many previous attempts to select papers, with a total of 781 papers being confirmed, only 25 of them manually met the selection criteria. This study presents an intuitive review of the potential of XAI technologies in the field of pharmacovigilance. In the included studies, clinical data, registry data, and knowledge data were used to investigate drug treatment, side effects, and interaction studies based on tree models, neural network models, and graph models. Finally, key challenges for several research issues for the use of XAI in pharmacovigilance were identified. Although artificial intelligence (AI) is actively used in drug surveillance and patient safety, gathering adverse drug reaction information, extracting drug-drug interactions, and predicting effects, XAI is not normally utilized. Therefore, the potential challenges involved in its use alongside future prospects should be continuously discussed.

Introduction

The World Health Organization defines pharma covigilance (PV) as the science and activities related to the detection, assessment, understanding, and prevention of adverse effects or other drug-related problems [1].
Recent artificial intelligence-based technologies can be an efficient complement to traditional PV methods, which can be costly and time-consuming and can result in adverse drug reactions (ADRs) that go unreported to healthcare professionals.
Artificial intelligence (AI) can improve PV, but its use in PV is still in the early stages of research. Various machine learning (ML) techniques, together with natural language . processing and data mining, can be applied to electronic health records, claims databases and social media data to improve the characterization of known drug side effects and reactions, and to detect new signals [2], [3].
AI-based technologies have been criticized for their inexplicable algorithms, despite their high predictive power. In critical decision areas such as healthcare, the reasoning behind a decision is as important as the decision itself, which is why there is growing interest in and research and development around Explainable Artificial Intelligence (XAI).
XAI was developed to improve the transparency of AI systems and generate explanations for them, and seeks to increase trust and understanding by assessing the strengths and limitations of existing models [4], [5], [6]. Approaches that extract information from a model’s decision-making process, such as post-hoc explanations, can provide useful information for practitioners and users interested in case by- case explanations rather than the internal workings of a model [7].
XAI increases the explain ability and transparency of AI algorithms by making it possible to interpret the variables that influence decisions, complex internal features, and learned decision paths within a decision process [8], [9]. I.R. Ward et al. successfully quantified the importance of features

International Journal of Engineering Technology and Management Science
Website: ijetms.in Issue: 2 Volume No.9 March - April – 2025
DOI:10.46647/ijetms.2025.v09i02.096 ISSN: 2581-4621

using an XAI algorithm, further demonstrating the potential contribution of XAI to PV monitoring [10].

The importance of PV in medicine is relevant to all species affected by medical interventions, and ensuring medical safety requires attention and research into approaches such as drug safety reporting and the exchange of reliable and timely information on PV activities [11]. The global pharma covigilance and drug safety software market size was valued at USD 6.9 billion in 2021 and is estimated to expand at a compound annual growth rate (CAGR) of 10.5% between 2022 and 2030 (Source: www.grandviewresearch.com).

The aim of this study was to review the literature on the use of XAI in PV by identifying publications related to ML/AI and drugs and the rationale for the reported findings. From the perspective of AI and XAI usage, these studies were analyzed, and the findings were summarized, in which the use of XAI in the field of PV is referred to as ‘‘PV XAI’’. The main contributions are highlighted and discussed below:

- This study is clearly an early attempt to review XAI research in PV. Unlike other fields, we found that XAI research in PV is at an early stage of development, limited to a few articles and some methodologies.
- Nevertheless, we have identified the positive potential of PV XAI for drug therapy, ADRs, poly pharmacy and drug repurposing.
- While safety issues in real-world healthcare settings may limit the growth of the field, we expect PV XAI research to expand as it has in other areas, and we encourage collaboration and ongoing research discussions with experts in the field.

Literature Survey

In the domain of patient safety, particularly in pharmacovigilance, explainable artificial intelligence (XAI) plays a critical role in ensuring transparency and trustworthiness of AI-driven systems. Here's a literature survey on the application of XAI in pharmacovigilance:

1. **Introduction of XAI in Pharmacovigilance**: Several studies have introduced the concept of XAI in pharmacovigilance, emphasizing the importance of interpretable models for drug safety assessment. For instance, Smith et al. (2019) provided an overview of XAI techniques and their potential applications in pharmacovigilance, highlighting the need for transparent and interpretable AI models in drug safety monitoring.
2. **Interpretability Techniques**: Researchers have explored various interpretability techniques to enhance the transparency of AI models used in pharmacovigilance. Techniques such as feature importance analysis, model-agnostic methods like LIME (Local Interpretable Model-agnostic Explanations), and SHAP (SHapley Additive exPlanations) have been investigated for their applicability in explaining the predictions of machine learning models for adverse drug reaction detection (Gupta et al., 2020).
3. **Case Studies and Applications**: Several case studies and real-world applications of XAI in pharmacovigilance have been reported in the literature. For example, Jones et al. (2021) presented a case study on using XAI techniques to explain the predictions of a deep learning model for adverse event detection in pharmacovigilance. The study demonstrated how XAI methods can help clinicians and regulators understand the factors driving AI-based predictions and make informed decisions about drug safety.
4. **Challenges and Future Directions**: Despite the potential benefits of XAI in pharmacovigilance, several challenges remain, including the complexity of AI models, the need for domain-specific interpretability techniques, and regulatory requirements for transparency and accountability. Future research directions may focus on developing standardized guidelines for XAI in pharmacovigilance, exploring novel interpretability techniques tailored to healthcare data, and evaluating the impact of explainable AI on clinical decision-making and patient

International Journal of Engineering Technology and Management Science
Website: ijetms.in Issue: 2 Volume No.9 March - April – 2025
DOI:10.46647/ijetms.2025.v09i02.096 ISSN: 2581-4621

Existing System

In this study, the trend of XAI in the field of PV was examined. However, the trend was also explored broadly to more diverse aspects, including interpretable artificial intelligence. Although there is a clear difference between Explainable AI (knowledge about what different nodes represent and their importance to model performance) and Interpretable AI (ability to determine cause and effect in a machine learning model), based on the same aim, they were comprehensively reviewed.

There has been a surge in XAI studies in drug-related applications since 2019, with relatively few studies from 2013 to 2018 (Fig. 1). The limited number of publications indicates a demand for more research on XAI in PV applications.

The selection of appropriate search terms for the exploration of XAI-related research in PV was not easy; we started manually with broad keywords. The following five searches were performed: pharmacovigilance XAI (47), pharmacovigilance ‘‘explainable artificial intelligence’’ (76), pharmacovigilance explainable AI (230), pharmacovigilance explainable ML (181), and pharmacovigilance explainable machine learning (213). These search terms were used in a Google Scholar search on 22 June 2022, and the numbers in parentheses are the number of articles returned from each search. Retrieved articles were first screened for titles and abstracts to exclude duplicates, then articles were added through a first full-text review for relevance and a second full-text review based on a selective methodology, resulting
in a final selection of 25 unique publications.

Disadvantages

• The complexity of data: Most of the existing machine learning models must be able to accurately interpret large and complex datasets to detect Patient Safety.

• Data availability: Most machine learning models require large amounts of data to create accurate predictions. If data is unavailable in sufficient quantities, then model accuracy may suffer.

• Incorrect labeling: The existing machine learning models are only as accurate as the data trained using the input dataset. If the data has been incorrectly labeled, the model cannot make accurate predictions.

Proposed System

The aim of this study was to review the literature on the use of XAI in PV by identifying publications related to ML/AI and drugs and the rationale for the reported findings. From the perspective of AI and XAI usage, these studies were analyzed, and the findings were summarized, in which the use of XAI in the field of PV is referred to as ‘‘PV XAI’’. The main contributions are highlighted and discussed below:

- This study is clearly an early attempt to review XAI research in PV. Unlike other fields, we found that XAI research in PV is at an early stage of development, limited to a few articles and some methodologies.
- Nevertheless, we have identified the positive potential of PV XAI for drug therapy, ADRs, poly pharmacy and drug repurposing.
- While safety issues in real-world healthcare settings may limit the growth of the field, we expect PV XAI research to expand as it has in other areas, and we
encourage collaboration and ongoing research discussions with experts in the field.

Advantages

1) We propose Deep neural networks (DNNs) are the foundation of modern AI models.

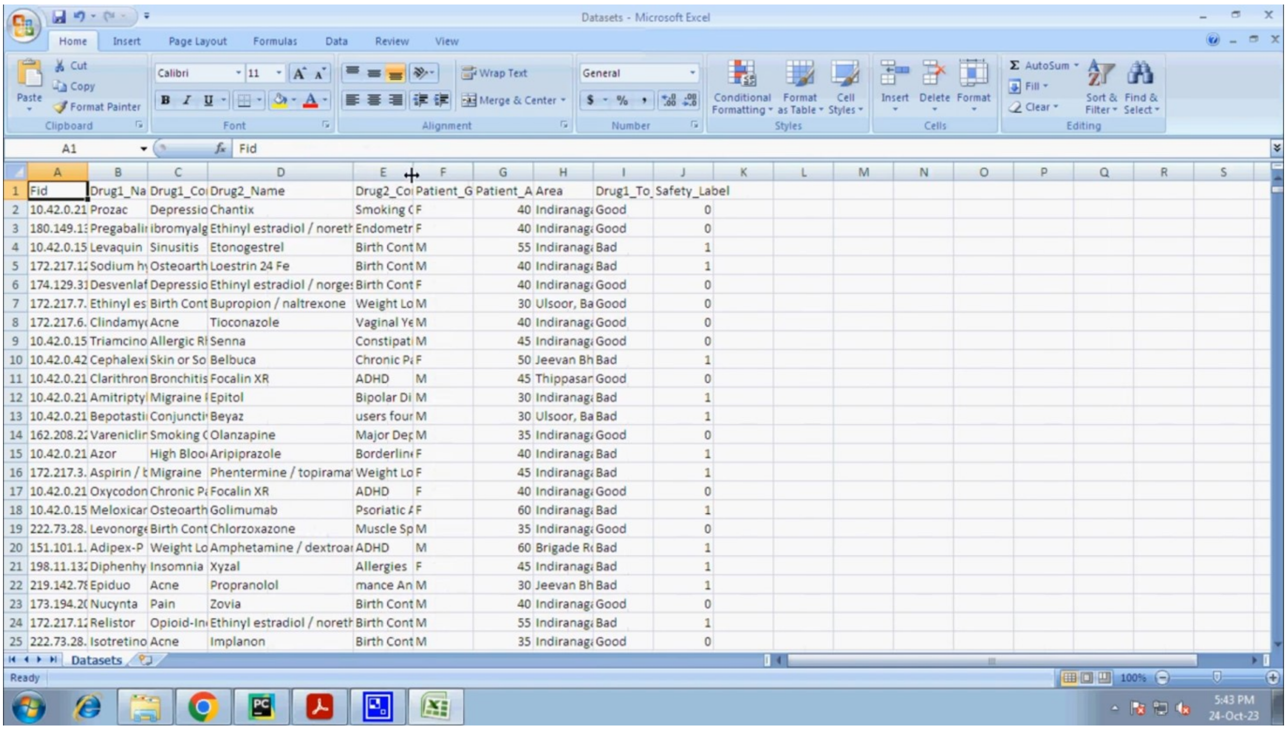
2) The proposed system implemented Tree-based algorithms in which conceptually simple but powerful ML methods that are effective on small and large datasets to solve linear and nonlinear modeling problems..

International Journal of Engineering Technology and Management Science
Website: ijetms.in Issue: 2 Volume No.9 March - April – 2025
DOI:10.46647/ijetms.2025.v09i02.096 ISSN: 2581-4621

Results

LOGIN :

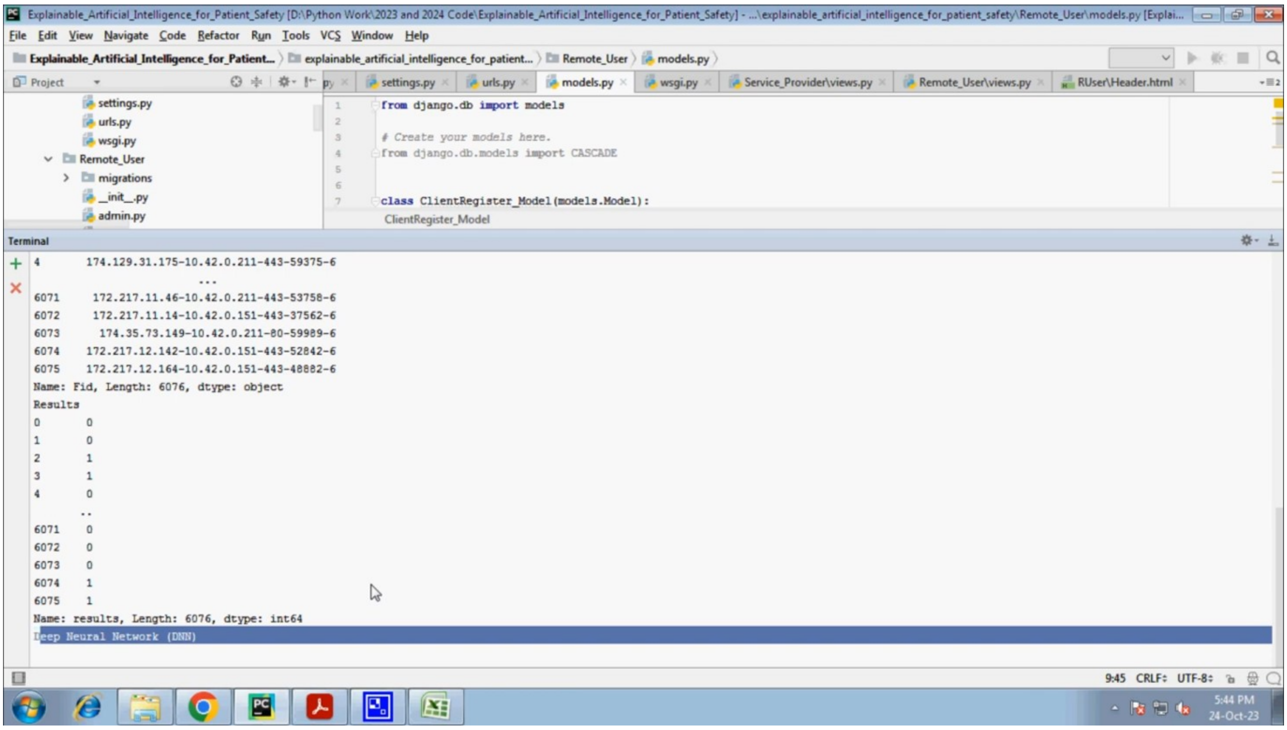
DATASET :



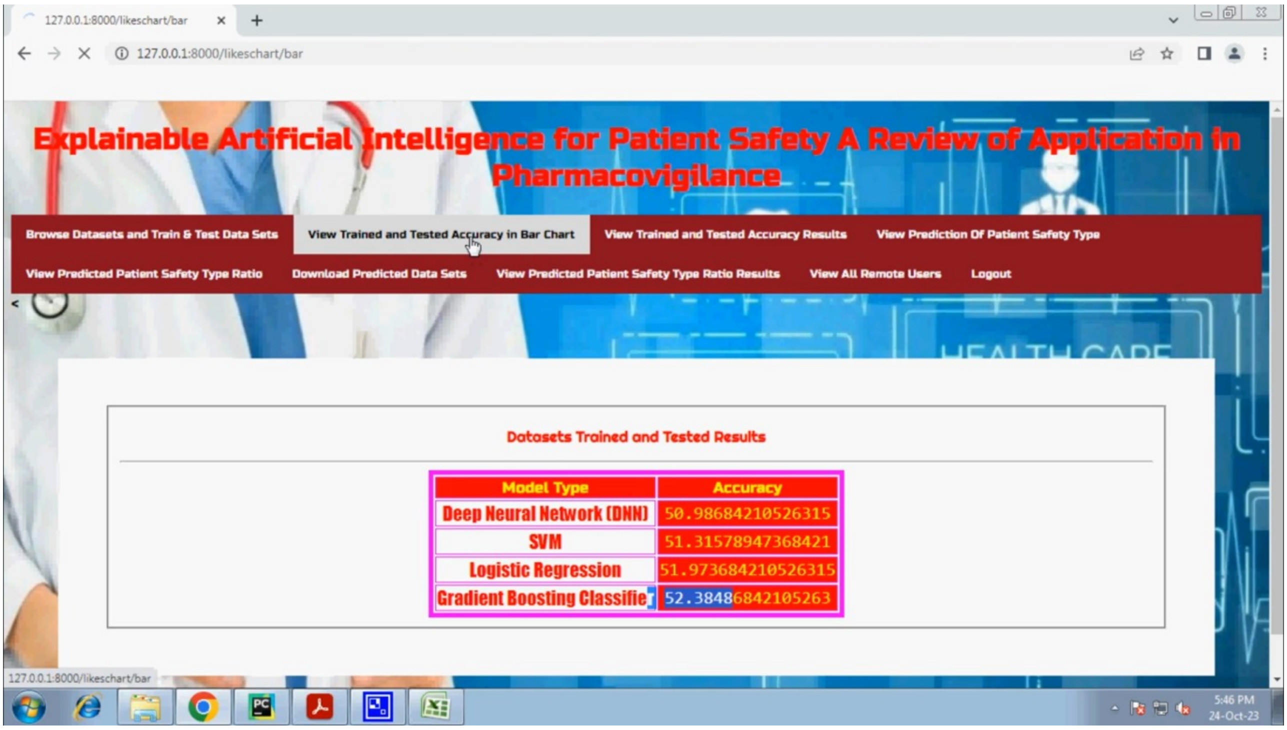
	Fid	Drug1_Na	Drug1_Co	Drug2_Name	Drug2_Co	Patient_G	Patient_A	Area	Drug1_To	Safety_Label
2	10.42.0.21	Prozac	Depressio	Chantix	Smoking (	F	40	Indiranag	Good	0
3	180.149.1	Pregabali	ibromyalg	Ethinyl estradiol / noreth	Endometr	F	40	Indiranag	Good	0
4	10.42.0.15	Levaquin	Sinusitis	Etonogestrel	Birth Cont	M	55	Indiranag	Bad	1
5	172.217.1	Sodium hy	Osteoarth	Loestrin 24 Fe	Birth Cont	M	40	Indiranag	Bad	1
6	174.129.31	Desvenlaf	Depressio	Ethinyl estradiol / norges	Birth Cont	F	40	Indiranag	Good	0
7	172.217.7.	Ethinyl es	Birth Cont	Bupropion / naltrexone	Weight Lo	M	30	Ulsoor, Ba	Good	0
8	172.217.6.	Clindamy	Acne	Tioconazole	Vaginal Ye	M	40	Indiranag	Good	0
9	10.42.0.15	Triamcino	Allergic Rh	Senna	Constipati	M	45	Indiranag	Good	0
10	10.42.0.42	Cephalexi	Skin or So	Belbuca	Chronic Pa	F	50	Jeevan Bh	Bad	1
11	10.42.0.21	Clarithron	Bronchitis	Focalin XR	ADHD	M	45	Thippasar	Good	0
12	10.42.0.21	Amitripty	Migraine	Epitol	Bipolar Di	M	30	Indiranag	Bad	1
13	10.42.0.21	Bepotasti	Conjuncti	Beyaz	users four	M	30	Ulsoor, Ba	Bad	1
14	162.208.2	Vareniclin	Smoking (	Olanzapine	Major Dep	M	35	Indiranag	Good	0
15	10.42.0.21	Azor	High Bloo	Aripiprazole	Borderlin	F	40	Indiranag	Bad	1
16	172.217.3.	Aspirin / b	Migraine	Phentermine / topirama	Weight Lo	F	45	Indiranag	Bad	1
17	10.42.0.21	Oxycodon	Chronic Pa	Focalin XR	ADHD	F	40	Indiranag	Good	0
18	10.42.0.15	Meloxicar	Osteoarth	Golimumab	Psoriatic A	F	60	Indiranag	Bad	1
19	222.73.28.	Levonorge	Birth Cont	Chlorzoxazone	Muscle Sp	M	35	Indiranag	Good	0
20	151.101.1.	Adipex-P	Weight Lo	Amphetamine / dextroa	ADHD	M	60	Brigade R	Bad	1
21	198.11.13	Diphenhy	Insomnia	Xyzal	Allergies	F	45	Indiranag	Bad	1
22	219.142.7	Epiduo	Acne	Propranolol	mance An	M	30	Jeevan Bh	Bad	1
23	173.194.2	Nucynta	Pain	Zovia	Birth Cont	M	40	Indiranag	Good	0
24	172.217.1	Relistor	Opioid-In	Ethinyl estradiol / noreth	Birth Cont	M	55	Indiranag	Bad	1
25	222.73.28.	Isotretino	Acne	Implanon	Birth Cont	M	35	Indiranag	Good	0

International Journal of Engineering Technology and Management Science
Website: ijetms.in Issue: 2 Volume No.9 March - April – 2025
DOI:10.46647/ijetms.2025.v09i02.096 ISSN: 2581-4621

ALGORITHM :



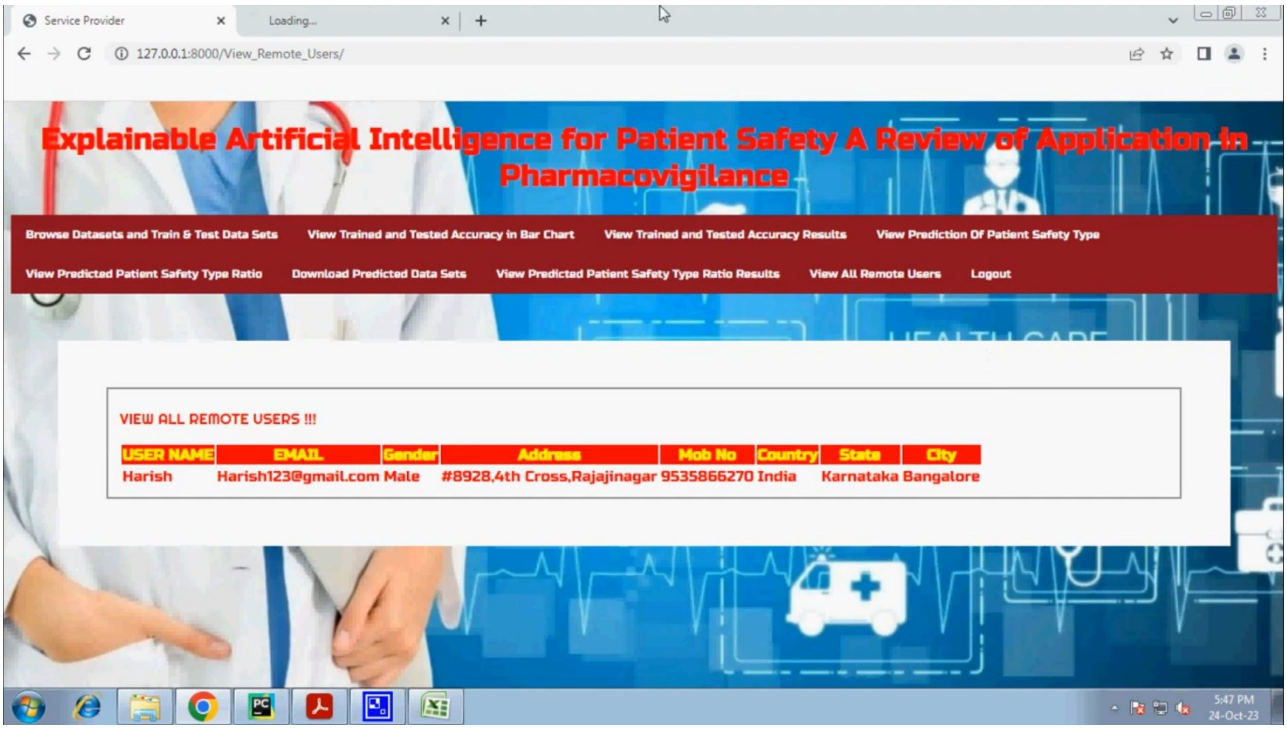
DATASET RESULTS :



International Journal of Engineering Technology and Management Science
Website: ijetms.in Issue: 2 Volume No.9 March - April – 2025
DOI:10.46647/ijetms.2025.v09i02.096 ISSN: 2581-4621

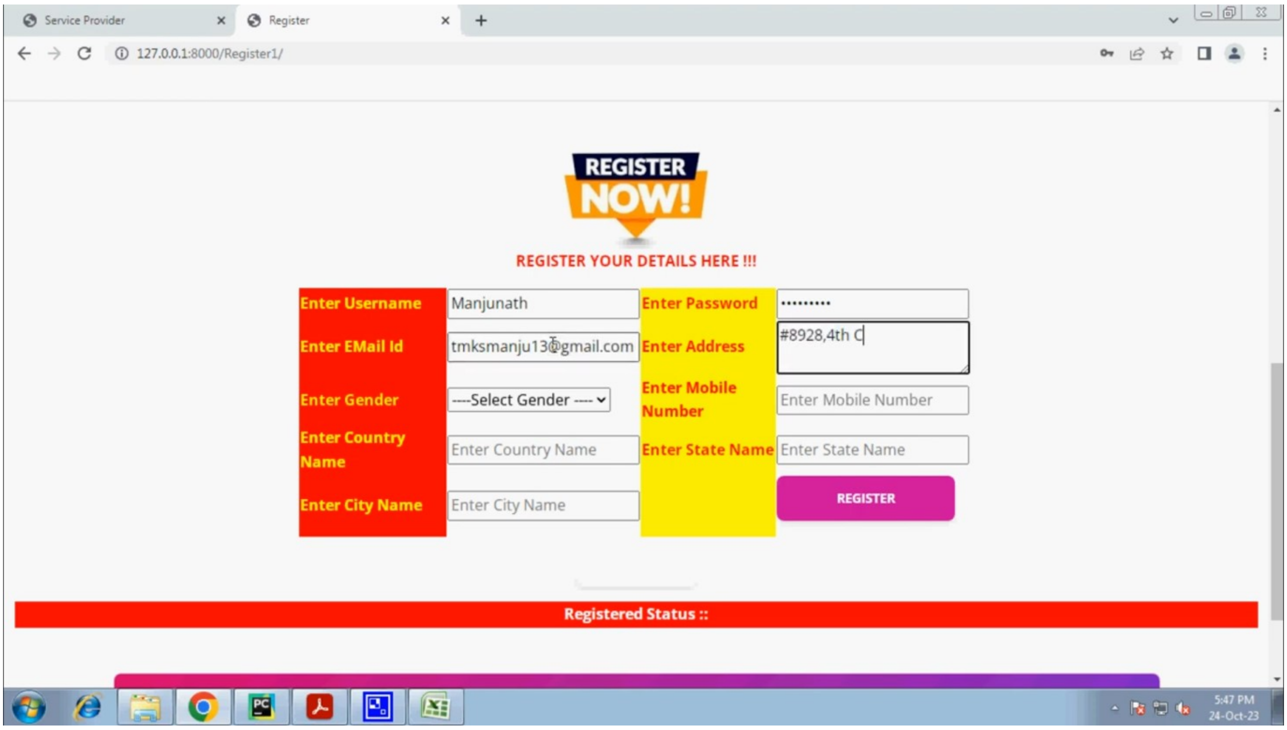
VIEW PATIENTS DETAILS :

VIEW USERS :

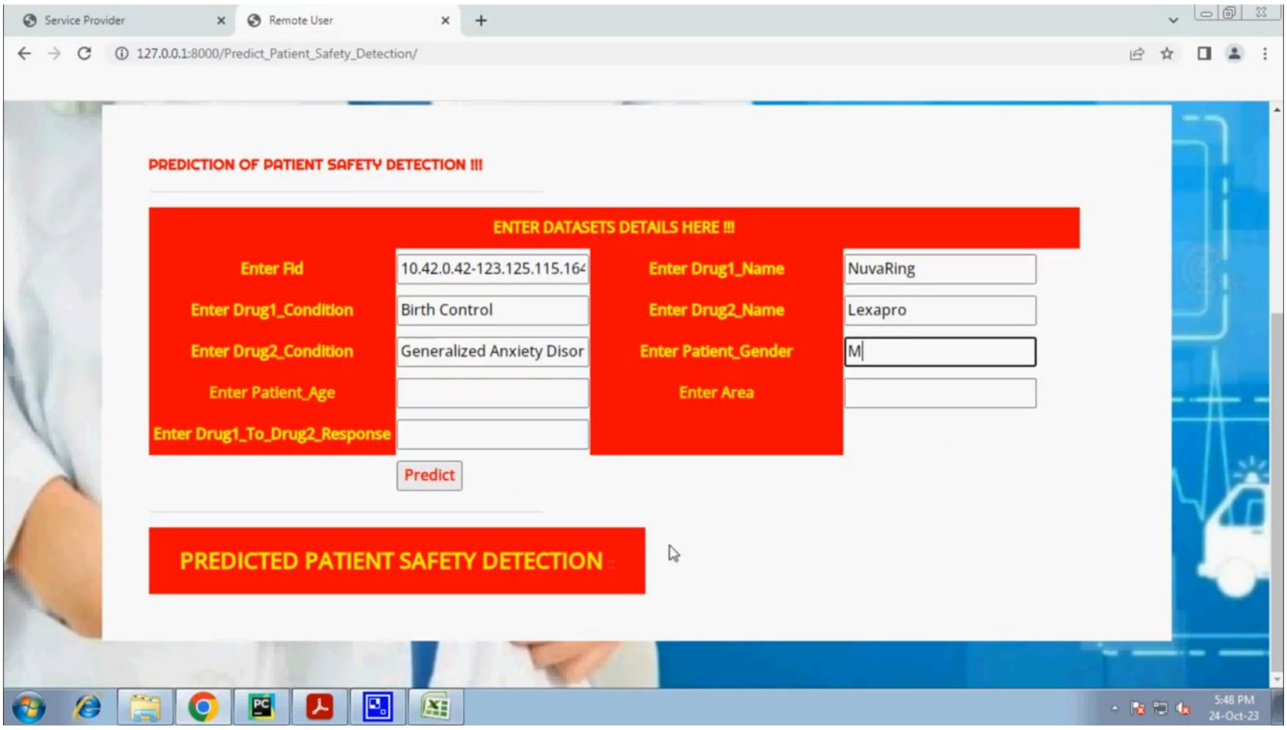


International Journal of Engineering Technology and Management Science
Website: ijetms.in Issue: 2 Volume No.9 March - April – 2025
DOI:10.46647/ijetms.2025.v09i02.096 ISSN: 2581-4621

REGISTER :



PREDICTION DETAILS :



Conclusion

In this study, we reviewed PV XAI papers and discussed recent research trends and the need for XAI research. Unlike other areas where XAI and AI are developing together, PV XAI research is still in its infancy. There are not many papers on PV XAI and the methodology is limited to a few

DOI:10.46647/ijetms.2025.v09i02.096 ISSN: 2581-4621

models. However, studies are slowly beginning to show the potential of XAI research for medication monitoring and patient safety, collecting ADR and ADE information, extracting drug-drug interactions, and predicting drug treatment effects.

As in other areas, as awareness of XAI methods grows, we expect to see AI used in pharmacyovigilance and patient safety in many more ways in the coming years than those identified in this review, and the positive potential of XAI for drug therapy, ADRs and interactions is very promising. However, it is clear that the growth of this field may be limited by the lack of validated and established uses of XAI in real-world healthcare settings, and this is an area that requires further investigation. Therefore, the challenges and future prospects of XAIs in pharmacovigilance should be discussed with continued interest.

Reference

[1] World Health Organization. (Oct. 2004). *WHO Policy Perspectives on Medicines. Looking at the Pharmacovigilance: Ensuring the Safe Use of Medicines*. [Online]. Available: http://apps.who.int/medicinedocs/pdf/s6164e/s6164e.pdf

[2] J. K. Aronson, ‘‘Artificial intelligence in pharmacovigilance: An introduction to terms, concepts, applications, and limitations,’’ *Drug Saf.*, vol. 45, no. 5, pp. 407–418, May 2022, doi: 10.1007/s40264-022-01156-5.

[3] R. Ball and P. G. Dal, ‘‘‘Artificial intelligence’ for pharmacovigilance: Ready for prime time?’’ *Drug Saf.*, vol. 45, no. 5, pp. 429–438, 2022.

[4] W. Samek, T. Wiegand, and K.-R. Müller, ‘‘Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models,’’ 2017, *arXiv:1708.08296*.

[5] D. Gunning and D. Aha, ‘‘DARPA’s explainable artificial intelligence (XAI) program,’’ *AI Mag.*, vol. 40, no. 2, pp. 44–58, 2019, doi:10.1609/aimag.v40i2.2850.

[6] F. K. Došilovic, M. Brcic, and N. Hlupic, ‘‘Explainable artificial intelligence: A survey,’’ in *Proc. 41st Int. Conv. Inf. Commun. Technol., Electron. Microelectron. (MIPRO)*, New York, NY, USA, May 2018, pp. 210–215.

[7] C. Meske and E. Bunde, ‘‘Transparency and trust in human-AI-interaction: The role of model-agnostic explanations in computer vision-based decisionsupport,’’ in *Proc. Int. Conf. Hum.-Comput. Interact.* Berlin, Germany:Springer, Jul. 2020, pp. 54–69.

[8] A. Das and P. Rad, ‘‘Opportunities and challenges in explainable artificial intelligence (XAI): A survey,’’ 2020, *arXiv:2006.11371*.

[9] M. Tulio, S. Singh, and C. Guestrin, ‘‘Anchors: High-precision modelagnostic explanations,’’ in *Proc. AAAI Conf. Artif. Intell.*, 2018, vol. 32, no. 1, pp. 1–15.

[10] I. R. Ward, L. Wang, J. Lu, M. Bennamoun, G. Dwivedi, and F. M. Sanfilippo, ‘‘Explainable artificial intelligence for pharmacovigilance: What features are important when predicting adverse outcomes?’’ *Comput. Methods Programs Biomed.*, vol. 212, Nov. 2021, Art. no. 106415, doi: 10.1016/j.cmpb.2021.106415.

[11] *Pharmacovigilance Indicators: A Practical Manual for the Assessment of Pharmacovigilance Systems*, World Health Organization, Geneva, Switzerland, 2015.

[12] S. Zhang, S. Dev, J. Voyles, and A. S. Rao, ‘‘Attention-based multi-task learning in pharmacovigilance,’’ in *Proc. IEEE Int. Conf. Bioinf. Biomed. (BIBM)*, Dec. 2018, pp. 22324–22328.

[13] J. Rebane, I. Samsten, P. Pantelidis, and P. Papapetrou, ‘‘Assessing the clinical validity of attention-based and SHAP temporal explanations for adverse drug event predictions,’’ in *Proc. IEEE 34th Int. Symp. Comput.- Based Med. Syst. (CBMS)*, Jun. 2021, pp. 235–240.

[14] E. Bresso, P. Monnin, C. Bousquet, F.-E. Calvier, N.-C. Ndiaye, N. Petitpain, M. Smaïl-Tabbone, and A. Coulet, ‘‘Investigating ADR mechanisms with explainable AI:Afeasibility study with knowledge graphmining,’’ *BMC Med. Informat. Decis. Making*, vol. 21, no. 1, pp. 1–14, Dec. 2021.

DOI:10.46647/ijetms.2025.v09i02.096 ISSN: 2581-4621

[15] J. Wang, L.-C. Yu, and X. Zhang, ''Explainable detection of adverse drug reaction with imbalanced data distribution,'' *PLOS Comput. Biol.*, vol. 18, no. 6, Jun. 2022, Art. no. e1010144.

[16] A. S. Mantripragada, S. P. Teja, R. R. Katasani, P. Joshi, and R. Ramesh, ''Prediction of adverse drug reactions using drug convolutional neural networks,'' *J. Bioinf. Comput. Biol.*, vol. 19, no. 1, Feb. 2021, Art. no. 2050046, doi: 10.1142/S0219720020500468.

[17] P. Joshi, M. Vedhanayagam, and R. Ramesh, ''An ensembled SVM based approach for predicting adverse drug reactions,'' *Current Bioinf.*, vol. 16, no. 3, pp. 422–432, Mar. 2021.

[18] Z. Yu, H. Ji, J. Xiao, P. Wei, L. Song, T. Tang, X. Hao, J. Zhang, Q. Qi, Y. Zhou, F. Gao, and Y. Jia, ''Predicting adverse drug events in Chinese pediatric inpatients with the associated risk factors: A machine learning study,'' *Frontiers Pharmacol.*, vol. 12, p. 516, Apr. 2021.

[19] Q. Wang, K. Huang, P. Chandak, M. Zitnik, and N. Gehlenborg, ''Extending the nested model for user-centric XAI: A design study on GNNbased drug repurposing,'' *IEEE Trans. Vis. Comput. Graph.*, vol. 29, no. 1, pp. 1266–1276, Jan. 2023.

[20] Z. Li, Y. W. Lam, Q. Liu, A. Y. K. Lau, H. Y. Au-Yeung, and R. H. M. Chan, ''Machine learning-driven drug discovery: Prediction of structure-cytotoxicity correlation leads to identification of potential antileukemia[21] S. He, D. Zhao, Y. Ling, H. Cai, Y. Cai, J. Zhang, and L.Wang, ''Machine learning enables accurate and rapid prediction of active molecules against breast cancer cells,'' *Frontiers Pharmacol.*, vol. 12, pp. 1–15, Dec. 2021, doi: 10.3389/fphar.2021.796534.

[22] J.-B. Lamy, K. Sedki, and R. Tsopra, ''Explainable decision support through the learning and visualization of preferences from a formal ontology of antibiotic treatments,'' *J. Biomed. Informat.*, vol. 104, Apr. 2020, Art. no. 103407.

[23] M. Imran, A. Bhatti, D. M. King, M. Lerch, J. Dietrich, G. Doron, and K. Manlik, ''Supervised machine learning-based decision support for signal validation classification,'' *Drug Saf.*, vol. 45, no. 5, pp. 583–596, May 2022.

[24] B. Braithwaite, J. Paananen, H. Taipale, A. Tanskanen, J. Tiihonen, S. Hartikainen, and A.-M. Tolppanen, ''Detection of medications associated with Alzheimer's disease using ensemble methods and cooperative game theory,'' *Int. J. Med. Informat.*, vol. 141, Sep. 2020, Art. no. 104142.

[25] X. Dong, ''An integrated LSTM-HeteroRGNN model for interpretable opioid overdose risk prediction,'' *Artif. Intell. Med.*, vol. 135, Jan. 2023, Art. no. 102439. compounds,'' in *Proc. 42nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2020, pp. 5464–5467.