



Machine Learning in Risk Stratification of Drug Events

Kiran Deshmukh *

Department of Biomedical Informatics, Global Institute for Clinical Research and Safety, Zurich, Switzerland

DESCRIPTION

Adverse Drug Events (ADEs) continue to pose a significant threat to patient safety and are among the leading causes of morbidity and hospital admissions worldwide. Despite ongoing efforts in pharmacovigilance, the complexity of drug reactions often influenced by multifactorial clinical, genetic, and behavioral factors has made early detection and risk stratification challenging. Recent advances in artificial intelligence, particularly Machine Learning (ML), offer a transformative approach to the identification, prediction, and stratification of drug-related risks by uncovering patterns in large, complex datasets that traditional statistical methods often fail to detect.

Machine learning enables the dynamic analysis of vast pharmacovigilance databases, including Electronic Health Records (EHRs), spontaneous reporting systems, and clinical trial data. Unlike rule-based systems, ML algorithms learn iteratively from data, adapting as new information becomes available. These models can identify subtle, non-linear relationships between variables, such as demographic features, drug combinations, comorbidities, and genetic predispositions, that might contribute to an adverse drug reaction. As such, machine learning provides a more individualized assessment of risk, aligning well with the goals of personalized medicine and proactive pharmacovigilance.

Several types of machine learning algorithms have been explored for risk stratification in pharmacovigilance. Supervised learning models, such as random forests, support vector machines, and gradient boosting classifiers, are trained on labeled datasets where outcomes of interest (e.g., occurrence of a drug event) are known. These models are particularly useful for building predictive tools that estimate the likelihood of adverse reactions in specific patient populations. Unsupervised learning techniques, such as clustering and dimensionality reduction, have also been applied to detect hidden patterns or rare adverse events in unlabeled data. These methods can group patients with similar risk profiles and detect outliers that may represent novel or unexpected reactions.

One of the notable applications of machine learning in this field is the automated signal detection in spontaneous adverse event reporting systems, such as the FDA's FAERS database or the WHO's Vigibase. ML algorithms can scan millions of reports to identify statistically significant associations between drugs and adverse events, often faster and more accurately than traditional disproportionality analyses. Incorporating Natural Language Processing (NLP) further enhances the ability of these models to analyze free-text clinical notes and case narratives, enriching the data available for prediction.

The real strength of machine learning lies in its potential to stratify risk not only retrospectively but also prospectively. In clinical settings, ML-powered decision support tools are being developed to alert healthcare providers in real time about patients at high risk of experiencing ADEs, based on current medication regimens, lab values, vital signs, and historical medical data. Such systems have shown promise in preventing events like drug-induced QT prolongation, nephrotoxicity, and bleeding complications, particularly in polypharmacy contexts where human oversight can be limited.

Despite the promise of ML in drug safety, challenges remain. Data quality and completeness are recurring issues, as missing or incorrect entries can bias model performance. Additionally, the "black box" nature of some algorithms raises concerns about transparency and interpretability, particularly in clinical decision-making where explanations of risk are essential. Regulatory agencies have begun to explore frameworks for validating and approving ML models in drug safety applications, emphasizing the need for transparency, reproducibility, and real-world performance assessment.

Moreover, ethical considerations such as algorithmic bias must be carefully managed. Training models on skewed datasets can inadvertently reinforce disparities in healthcare, disproportionately impacting underrepresented populations. It is therefore essential to develop models that are not only accurate but also fair and generalizable across diverse patient groups.

Collaborative initiatives between academia, regulatory bodies,

Correspondence to: Kiran Deshmukh, Department of Biomedical Informatics, Global Institute for Clinical Research and Safety, Zurich, Switzerland, E-mail: k.deshmukh@gicrs-zurich.edu

Received: 03-Mar-2025, Manuscript No. JP-25-28633; **Editor assigned:** 05-Mar-2025, PreQC No. JP-25-28633 (PQ); **Reviewed:** 19-Mar-2025, QC No. JP-25-28633; **Revised:** 26-Mar-2025, Manuscript No. JP-25-28633 (R); **Published:** 02-Apr-2025, DOI: 10.35248/2329-6887.25.13.501

Citation: Deshmukh K (2025). Machine Learning in Risk Stratification of Drug Events. J Pharmacovigil. 13:501.

Copyright: © 2025 Deshmukh K. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

and the pharmaceutical industry are currently underway to standardize methodologies and share best practices for ML in pharmacovigilance. Open-access data repositories and model libraries are helping accelerate research while promoting

transparency and reproducibility. The future may also see integration of genomics, wearable device data, and real-world evidence into ML frameworks, further enriching their predictive power.