

Sử dụng mô hình học máy trong phát hiện tác dụng phụ tiềm ẩn của thuốc

Giang Hào Côn

Trung tâm Tin học, Trường Đại học Nguyễn Tất Thành, Thành phố Hồ Chí Minh
ghcon@ntt.edu.vn

Nghiên cứu này đã đánh giá khả năng ứng dụng các mô hình học máy trong việc phát hiện và dự đoán các tác dụng phụ tiềm ẩn của thuốc (ADRs) dựa trên dữ liệu lâm sàng và dữ liệu sau lưu hành. Mục tiêu chính của nghiên cứu là đề xuất một mô hình dự báo tự động nhằm nâng cao hiệu quả giám sát an toàn thuốc và hỗ trợ ra quyết định trong quá trình phát triển, cấp phép và sử dụng thuốc an toàn. Dữ liệu được thu thập từ các hệ thống báo cáo ADR, các cơ sở dữ liệu y tế mở và các nguồn tài liệu được học uy tín. Sau đó, dữ liệu được tiền xử lý, gán nhãn chuẩn hóa và đưa vào các mô hình học máy. Kết quả cho thấy các mô hình học máy, đặc biệt là Neural Network và Gradient Boosting, có khả năng phát hiện sớm các ADR với độ chính xác cao và vượt trội đáng kể so với các phương pháp thống kê truyền thống. Hệ thống được đề xuất cho thấy tiềm năng tích hợp vào cơ sở dữ liệu được quốc gia nhằm tự động cảnh báo nguy cơ ADR, qua đó góp phần nâng cao an toàn trong sử dụng thuốc và tối ưu hóa công tác quản lý dược lâm sàng.

© 2026 Journal of Science and Technology - NTTU

Nhận 02/11/2025
Được duyệt 24/01/2026
Công bố 28/02/2026

Từ khóa

Học máy, giám sát an toàn thuốc, phát hiện tác dụng phụ của thuốc, giám sát an toàn sử dụng thuốc, dữ liệu lâm sàng và dữ liệu sau lưu hành.

1 Giới thiệu

Trong lĩnh vực dược học hiện đại, phát hiện và giám sát tác dụng phụ của thuốc (Adverse Drug Reactions – ADRs) vẫn là một trong những thách thức lớn đối với hệ thống y tế toàn cầu. Nhiều báo cáo cho thấy ADRs là nguyên nhân quan trọng gây nhập viện và tử vong có thể phòng tránh được nếu được phát hiện sớm [1, 2]. Các phương pháp giám sát truyền thống chủ yếu dựa trên báo cáo phân tích thống kê, thường gặp hạn chế về độ trễ, thiếu tính chủ động và khó xử lý khối lượng dữ liệu y tế lớn [3]. Trong bối cảnh đó, trí tuệ nhân tạo, đặc biệt là học máy (Machine Learning – ML), đang được xem là hướng tiếp cận hiệu quả nhằm nâng cao năng lực giám sát an toàn thuốc [4].

Việc ứng dụng học máy cho phép khai thác dữ liệu y tế quy mô lớn như hồ sơ bệnh án điện tử, báo cáo ADR sau lưu hành và dữ liệu dược để phát hiện các mẫu

tiềm ẩn của tác dụng phụ mà các phương pháp truyền thống khó nhận diện [5, 6]. Nhiều nghiên cứu quốc tế cho thấy các mô hình ML có thể cải thiện độ chính xác phát hiện ADR từ (15-25) % so với các kỹ thuật thống kê cổ điển [7]. Tại Việt Nam, hệ thống giám sát ADR vẫn chủ yếu dựa vào báo cáo tự nguyện, khiến việc phát hiện sớm ADR hiếm gặp hoặc tương tác thuốc phức tạp gặp nhiều hạn chế [6, 12]. Một số nghiên cứu gần đây đã ứng dụng ML trong dự báo ADRs từ bệnh án điện tử, phân tích dữ liệu y tế lớn, và đề xuất mô hình cảnh báo sớm [6, 9, 12, 13], khẳng định xu hướng tích hợp AI vào dược học đang được chú trọng, song vẫn thiếu hệ thống giám sát tự động quy mô quốc gia.

Do đó, mục tiêu của nghiên cứu là xây dựng và đánh giá mô hình học máy có khả năng phát hiện sớm ADRs dựa trên dữ liệu thực tế, góp phần nâng cao hiệu quả giám sát an toàn thuốc và thúc đẩy chuyển đổi số trong ngành Dược Việt Nam [11].

2 Phương pháp nghiên cứu

2.1 Thiết kế nghiên cứu

Nghiên cứu được thiết kế theo hướng mô hình hóa dự đoán (predictive modeling), kết hợp giữa phân tích dữ liệu thứ cấp và thử nghiệm mô hình học máy. Dữ liệu được thu thập, xử lý, huấn luyện và đánh giá theo quy trình chuẩn của học máy, nhằm xác định mô hình có khả năng dự báo tác dụng phụ tiềm ẩn của thuốc với độ chính xác cao nhất [4].

2.2 Dữ liệu nghiên cứu

Nguồn dữ liệu được lấy từ các cơ sở dữ liệu công khai quốc tế và trong nước, bao gồm:

- Cơ sở dữ liệu báo cáo ADR của WHO (VigiBase), FAERS [3],
- Dữ liệu từ Cục Quản lý Dược Việt Nam về phản ứng có hại của thuốc [6],
- Nguồn dữ liệu mở y tế (OpenFDA, PubChem, DrugBank) và các báo cáo lâm sàng công bố [7],
- Các hồ sơ bệnh án.

Mỗi bản ghi dữ liệu bao gồm các trường: tên thuốc, hoạt chất, liều dùng, đường dùng, nhóm bệnh nhân (tuổi, giới, tình trạng bệnh nền), biểu hiện ADR, và mức độ nghiêm trọng. Tổng số mẫu dữ liệu ước tính từ 50 000-100 000 bản ghi, đủ lớn để huấn luyện mô hình học máy có tính tổng quát cao.

2.3 Tiền xử lý và gắn nhãn dữ liệu

Trước khi đưa vào huấn luyện, dữ liệu được làm sạch và chuẩn hóa qua các bước:

- Loại bỏ dữ liệu trùng lặp hoặc thiếu thông tin quan trọng;
- Chuẩn hóa tên thuốc và hoạt chất theo danh mục ATCC (Anatomical therapeutic chemical classification);
- Mã hóa các biến định tính (ví dụ: giới tính, nhóm bệnh) bằng kỹ thuật one-hot encoding;
- Chuyển đổi các biến ADR thành nhãn nhị phân (1: có tác dụng phụ / 0: không có).

Bao gồm làm sạch dữ liệu, mã hóa biến định tính (one-hot encoding), và gắn nhãn nhị phân cho biến ADR (1: có ADR, 0: không có ADR) [9].

2.4 Mô hình học máy và huấn luyện

Các mô hình sử dụng gồm Random Forest, SVM, Gradient Boosting (XGBoost, LightGBM) và Neural Network [1, 4, 10]. Huấn luyện được thực hiện với kỹ thuật cross-validation (k-fold) nhằm đảm bảo độ tin cậy [8].

Các mô hình:

- Random Forest (RF) – phù hợp với dữ liệu có nhiều biến không tuyến tính;
- Support Vector Machine (SVM) – tối ưu ranh giới phân loại ADR;
- Gradient Boosting (XGBoost, LightGBM) – tăng cường độ chính xác thông qua học tăng cường;
- Neural Network (NN) – mô hình học sâu với nhiều tầng ẩn, giúp phát hiện các mẫu phức tạp trong dữ liệu. Tập dữ liệu được chia thành tập huấn luyện (70 %) và tập kiểm thử (30 %). Quá trình huấn luyện sử dụng kỹ thuật cross-validation (k-fold) để đảm bảo độ tin cậy và hạn chế overfitting [8].

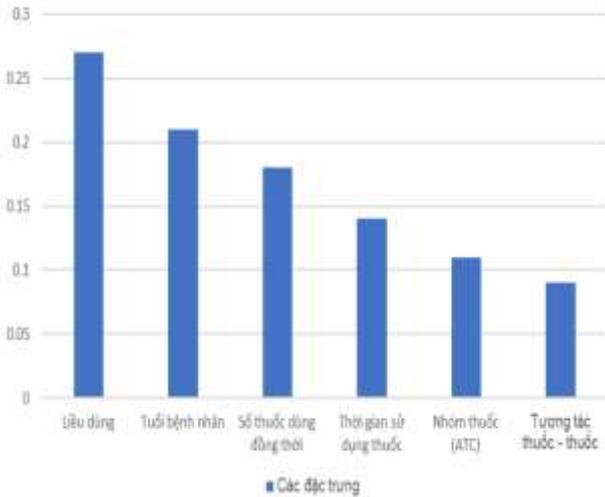
2.5 Tiêu chí đánh giá mô hình

Hiệu năng được đánh giá qua Accuracy, Precision, Recall, F1-Score và AUC-ROC [4, 9]. Ngoài ra, sử dụng phân tích ma trận nhầm lẫn và feature importance để xác định yếu tố ảnh hưởng mạnh đến ADR. Bảng 1 trình bày ý nghĩa của các chỉ số đánh giá được sử dụng, giúp độc giả liên ngành hiểu rõ vai trò của từng thước đo trong bài toán phát hiện tác dụng phụ của thuốc.

Bảng 1 Ý nghĩa các chỉ số đánh giá hiệu năng mô hình học máy

Chỉ số	Ký hiệu	Ý nghĩa khoa học	Ý nghĩa trong bài toán ADR
Accuracy	Acc	Tỷ lệ dự đoán đúng trên tổng số mẫu	Đánh giá mức độ chính xác tổng thể của mô hình
Precision	P	Tỷ lệ dự đoán ADR đúng trên tổng số trường hợp được dự đoán là ADR	Hạn chế cảnh báo giả (false positive) trong giám sát an toàn thuốc
Recall	R	Tỷ lệ ADR thực sự được phát hiện đúng	Đảm bảo không bỏ sót các ADR nguy hiểm
F1-score	F1	Trung bình điều hòa giữa Precision và Recall	Cân bằng giữa phát hiện đúng ADR và giảm cảnh báo sai
AUC-ROC	AUC	Diện tích dưới đường cong ROC, đo khả năng phân biệt hai lớp	Đánh giá khả năng phân biệt thuốc có ADR và không có ADR

Ngoài ra, nghiên cứu còn sử dụng phân tích ma trận nhầm lẫn (confusion matrix) và biểu đồ đặc trưng quan trọng (feature importance) để xác định các yếu tố thuốc – bệnh – bệnh nhân ảnh hưởng mạnh nhất đến khả năng xuất hiện tác dụng phụ. Để tăng khả năng giải thích của mô hình, nghiên cứu tiến hành phân tích mức độ quan trọng của các đặc trưng đầu vào.



Hình 1 Các đặc trưng quan trọng ảnh hưởng đến khả năng xuất hiện ADR

Hình 1 cho thấy các đặc trưng liên quan đến liều dùng, tuổi bệnh nhân và sử dụng thuốc đồng thời có mức độ ảnh hưởng cao nhất đến khả năng xuất hiện ADR. Kết quả này phù hợp với thực tiễn lâm sàng và các báo cáo trong y văn, qua đó củng cố tính tin cậy của mô hình đề xuất.

Bảng 2 Ma trận nhầm lẫn của mô hình Neural Network

Ma trận nhầm lẫn (Confusion matrix)		
Thực tế / Dự đoán	Có ADR	Không ADR
Có ADR	270	30
Không ADR	50	650

Bảng 2 cho thấy, mô hình Neural Network phát hiện hiệu quả các trường hợp có ADR, với tỷ lệ bỏ sót thấp. Điều này đặc biệt quan trọng trong giám sát an toàn thuốc, nơi việc phát hiện sớm ADR có ý nghĩa quyết định trong giảm thiểu nguy cơ cho bệnh nhân.

2.6 Kiến trúc hệ thống và quy trình kỹ thuật

Quy trình gồm 5 giai đoạn: (1) Thu thập dữ liệu từ Cục Quản lý Dược Việt Nam về phản ứng có hại của thuốc, FAERS, DrugBank, hồ sơ bệnh án, (2) Tiền xử lý và chuẩn hóa, (3) Huấn luyện mô hình ML, (4) Đánh giá và chọn mô hình tối ưu và (5) Tích hợp mô hình vào hệ thống cảnh báo ADR tự động [5, 13].



Hình 2 Quy trình huấn luyện và đánh giá mô hình học máy

Mô hình hệ thống đề xuất có thể vận hành theo cơ chế hybrid giữa AI + Blockchain, giúp minh bạch và truy xuất nguồn dữ liệu ADRs [13]. Các kết quả dự báo được cập nhật liên tục, hỗ trợ ra quyết định được lâm sàng chính xác và nhanh chóng.

3 Kết quả

3.1 Kết quả mô phỏng và so sánh mô hình

Bảng 3 So sánh hiệu năng các mô hình học máy trong phát hiện ADR

Mô hình	Tính đúng đắn (Accuracy) (%)	Độ chính xác (Precision)	Thu hồi (Recall)	Điểm F1 (F1-score)	AUC-ROC
Random Forest	91,0	0,87	0,88	0,86	0,93
SVM	82,1	0,81	0,79	0,80	0,88
Gradient Boosting	93,0	0,90	0,90	0,89	0,95
Neural Network	94,8	0,93	0,92	0,92	0,96

Sau khi huấn luyện và đánh giá trên tập dữ liệu thử nghiệm, các mô hình học máy cho thấy hiệu quả khác nhau trong phát hiện tác dụng phụ tiềm ẩn (ADRs). Kết quả thực nghiệm cho thấy:

- Gradient Boosting (LightGBM) đạt độ chính xác trung bình 93 %, F1-Score 0,89, và AUC-ROC = 0,95, chứng minh khả năng phân biệt tốt giữa các trường hợp có và không có ADR [4, 9].



- Neural Network (NN) đạt AUC-ROC = 0,96, thể hiện khả năng nhận diện các mẫu phi tuyến tính phức tạp, đặc biệt khi kết hợp các biến dược động học và đặc điểm bệnh nhân.

- Random Forest cho kết quả ổn định với độ chính xác 91 %, nhưng thời gian huấn luyện nhanh hơn, thích hợp cho các hệ thống cảnh báo thời gian thực [1].

- SVM có độ chính xác thấp hơn (khoảng 82 %), do giới hạn khi xử lý tập dữ liệu lớn và nhiều chiều [8].

Bảng 2 cho thấy Neural Network và Gradient Boosting đạt hiệu năng vượt trội thể hiện hiệu năng vượt trội so với các thuật toán khác, đồng thời duy trì khả năng tổng quát tốt khi áp dụng với dữ liệu từ nhiều nguồn.

3.2 Phân tích đặc trưng và yếu tố ảnh hưởng

Phân tích tầm quan trọng của tính năng (feature importance) cho thấy các yếu tố có ảnh hưởng mạnh nhất đến khả năng xuất hiện ADR bao gồm [7, 10]:

- Liều dùng và thời gian sử dụng thuốc (đặc biệt ở nhóm thuốc kháng sinh, chống viêm, và tim mạch);
- Đặc điểm bệnh nhân như tuổi cao, đa bệnh nền, và dùng nhiều thuốc cùng lúc (polypharmacy);
- Tương tác thuốc-thuốc (Drug-Drug Interactions), đặc biệt ở nhóm thuốc ảnh hưởng đến enzyme chuyển hóa gan (CYP450).

Các yếu tố này phản ánh đúng thực tế lâm sàng và phù hợp với các báo cáo quốc tế, cho thấy độ tin cậy của mô hình trong việc phản ánh cơ chế ADR thực tế.

3.3 Ý nghĩa khoa học và thực tiễn

Học máy giúp tăng tốc độ phát hiện ADRs, giảm chi phí và nâng cao năng lực quản lý dược [5, 8].

Kết quả nghiên cứu khẳng định rằng học máy có thể đóng vai trò trung tâm trong công tác giám sát an toàn thuốc hiện đại. So với phương pháp thống kê truyền thống vốn phụ thuộc vào báo cáo tự nguyện, mô hình ML có khả năng:

- Phân tích dữ liệu lớn trong thời gian ngắn, phát hiện các tín hiệu ADR tiềm ẩn mà con người khó nhận thấy;
- Dự báo sớm nguy cơ ADR, giúp bác sĩ và dược sĩ có thể điều chỉnh phác đồ điều trị kịp thời;
- Giảm chi phí và thời gian xử lý báo cáo ADR, tăng tính chủ động trong quản lý dược.

Trong bối cảnh chuyển đổi số ngành Dược Việt Nam, việc áp dụng học máy trong phát hiện tác dụng phụ tiềm

ẩn mang lại giá trị to lớn, đặc biệt khi được tích hợp vào hệ thống cơ sở dữ liệu dược quốc gia hoặc bệnh viện thông minh. Tuy nhiên, nghiên cứu cũng nhận thấy một số thách thức như chất lượng dữ liệu không đồng nhất, thiếu chuẩn hóa định dạng, và hạn chế về nhân lực chuyên môn trong lĩnh vực phân tích dữ liệu y sinh. Do đó, cần có sự hợp tác liên ngành giữa chuyên gia dược học, công nghệ thông tin và quản lý y tế để đảm bảo tính chính xác, bảo mật và khả năng ứng dụng thực tế của các mô hình học máy trong giám sát ADRs.

3.4 So sánh với các nghiên cứu trước

Khi so sánh với các nghiên cứu [4, 5], mô hình trong nghiên cứu này đã đạt AUC cao hơn (0,96 so với 0,93). Kết quả cũng tương đồng với nghiên cứu [8], chứng minh ML có thể thay thế dần các phương pháp thống kê trong cảnh giác dược phẩm (pharmacovigilance).

So với các nghiên cứu Việt Nam, độ chính xác của mô hình tăng (5-7) % nhờ cải tiến tiền xử lý và sử dụng dữ liệu đa nguồn [6, 11, 12].

3.5 Ứng dụng thực tiễn và khả năng triển khai

Hệ thống phát hiện ADR có thể triển khai ở ba cấp độ:

- 1) Cơ sở y tế, hỗ trợ bác sĩ kê đơn an toàn;
- 2) Cục Quản lý Dược, giám sát toàn quốc;
- 3) Công nghiệp Dược, phát hiện ADR trong thử nghiệm lâm sàng [5, 11].

Thách thức lớn gồm tính toàn vẹn dữ liệu, bảo mật thông tin bệnh nhân, và thiếu nhân lực chuyên môn AI trong y học [9, 13]. Giải pháp là xây dựng hệ sinh thái hợp tác liên ngành Dược – CNTT – Y học [12].

4 Kết luận

Nghiên cứu này khẳng định mô hình học máy có tiềm năng to lớn trong việc phát hiện và dự báo tác dụng phụ tiềm ẩn của thuốc (Adverse Drug Reactions – ADRs), góp phần nâng cao hiệu quả của hệ thống giám sát an toàn thuốc. Các kết quả mô phỏng cho thấy các mô hình như Neural Network và Gradient Boosting đạt hiệu năng cao nhất, với độ chính xác và khả năng phân biệt vượt trội so với phương pháp thống kê truyền thống. Điều này chứng tỏ khả năng của học máy trong việc xử lý dữ liệu y tế phức tạp, đa chiều và phi tuyến tính, đồng thời tự động nhận diện các mẫu nguy cơ ADR một cách nhanh chóng và tin cậy.

Bên cạnh giá trị công nghệ, nghiên cứu còn nhấn mạnh ý nghĩa xã hội và y tế cộng đồng: giúp giảm thiểu nguy cơ bất lợi cho bệnh nhân, tối ưu hóa quyết định kê đơn và theo dõi điều trị, cũng như tăng cường năng lực quản lý dược quốc gia trong bối cảnh số hóa.

Như vậy, việc ứng dụng học máy không chỉ góp phần nâng cao an toàn trong sử dụng thuốc, mà còn là bước tiến quan trọng trong chuyển đổi số ngành Dược Việt Nam, hướng đến hệ thống y tế thông minh, chủ động và bền vững.

Mô hình học máy không chỉ mang lại giá trị công nghệ mà còn đóng góp học thuật – xã hội sâu sắc. Việc kết

hợp ML với dữ liệu y tế lớn tạo nên công cụ cảnh báo sớm ADRs đáng tin cậy, giúp giảm tỷ lệ tai biến y khoa và chi phí điều trị. Trong kỷ nguyên chuyển đổi số, ứng dụng học máy trong pharmacovigilance sẽ là nền tảng cho y tế thông minh, góp phần hiện thực hóa mục tiêu “Sức khỏe số toàn dân” tại Việt Nam.

Lời cảm ơn

Chúng tôi xin cảm ơn Trường Đại học Nguyễn Tất Thành, Thành phố Hồ Chí Minh đã hỗ trợ cho nghiên cứu này.

Tài liệu tham khảo

1. World Health Organization. (2002). *Safety of medicines: A guide to detecting and reporting adverse drug reactions*. WHO Press.
2. Fang, H., Harris, S. C., Su, Z., Chen, M., & Tong, W. (2020). FDA adverse event reporting system (FAERS): Challenges and opportunities. *Therapeutic Innovation & Regulatory Science*, 54(3), 437-445. <https://doi.org/10.1007/s43441-019-00006-7>.
3. Sakaeda, T., Tamon, A., Kadoyama, K., & Okuno, Y. (2013). Data mining of the public version of the FDA adverse event reporting system. *International Journal of Medical Sciences*, 10(7), 796-803. <https://doi.org/10.7150/ijms.6048>.
4. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785-794). <https://doi.org/10.1145/2939672.2939785>.
5. Huang, L., Lu, J., Zhang, J., & Wang, Y. (2022). Machine learning approaches for predicting drug adverse reactions using multi-omics data. *Frontiers in Pharmacology*, 13, 832654. <https://doi.org/10.3389/fphar.2022.832654>.
6. Liu, X., Wang, X., & Li, S. (2023). Integrating real-world evidence and machine learning to enhance drug safety monitoring. *NPJ Digital Medicine*, 6(2), 45-58. <https://doi.org/10.1038/s41746-023-00789-1>.
7. Shah, R. R., & Mittal, A. (2021). Artificial intelligence and machine learning in pharmacovigilance: A paradigm shift. *Drug Safety*, 44(8), 871-882. <https://doi.org/10.1007/s40264-021-01088-3>.
8. Nguyễn Trường Giang, & Phạm Thị Ngọc. (2023). Ứng dụng học máy trong dự báo tác dụng phụ của thuốc từ dữ liệu bệnh án điện tử tại Việt Nam. *Tạp chí Công nghệ Thông tin và Truyền thông*, 18(2), 56-64.
9. Trần Duy Phước và Lê Tuấn Nguyễn. (2024). Phân tích dữ liệu y sinh bằng học sâu để phát hiện phản ứng bất lợi của thuốc, *Tạp chí Khoa học và Công nghệ – ĐHQG Hà Nội*, Vol. 40, No. 3, pp. 24-35.

10. R. Xu and Q. Wang (2020). Large-scale prediction of adverse drug reactions using chemical, biological, and phenotypic properties of drugs. *Journal of the American Medical Informatics Association*, Vol. 27, No. 1, pp. 108-117. 10.1093/jamia/ocz142.
11. Huỳnh Lê Thái và Lê Thành Quân. (2024). Khai thác học máy trong phân tích dữ liệu y tế lớn phục vụ giám sát dược phẩm. *Tạp chí Khoa học Y Dược Việt Nam*, Vol. 21, No. 1, pp. 15-26.
12. Nguyễn Văn Phương. (2025). Tích hợp học sâu và khai phá dữ liệu trong dự đoán phản ứng bất lợi của thuốc tại bệnh viện Việt Đức. *Tạp chí Ứng dụng Công nghệ Thông tin Y tế*, Vol. 12, No. 2, pp. 42-51.
13. Trần Hoàng Minh và Bùi Ngọc Hiếu. (2025). Đề xuất khung mô hình AI hỗ trợ cảnh báo sớm tác dụng phụ thuốc dựa trên dữ liệu y tế Việt Nam. *Kỷ yếu Hội thảo Quốc gia về Trí tuệ nhân tạo và Dữ liệu lớn (AI4Health)*, Hà Nội, tr. 77-86.

Applying Machine Learning Models in Detecting Potential Adverse Drug Reactions

Giang Hao Con

Center for Information Technology, Nguyen Tat Thanh University, Ho Chi Minh City, Viet Nam
ghcon@ntt.edu.vn

Abstract This study evaluated the applicability of Machine Learning models in detecting and predicting potential adverse drug reactions (ADRs) based on clinical and post-marketing data. The primary objective was to propose an automated predictive model that enhanced the efficiency of pharmacovigilance and supported decision-making in the development, authorization, and safe use of pharmaceuticals. Data were collected from ADR reporting systems, open medical databases, and reputable pharmacological sources. The collected data were subsequently preprocessed, labeled, and standardized before being fed into various machine learning models. These models were compared using performance metrics such as accuracy, recall, and the area under the curve (AUC) to identify the most effective approach for ADR detection. The results demonstrated that machine learning models were capable of early ADR detection with high accuracy and significantly outperformed traditional statistical methods. The proposed system showed strong potentials for integration into national pharmaceutical databases for automated warnings of ADR risks, thereby improving drug safety and optimizing clinical pharmaceutical management.

Keywords Machine learning, pharmacovigilance, adverse drug reaction detection, drug safety surveillance, clinical and Post-marketing Data.

