British Journal of Computer, Networking and Information Technology
ISSN: 2689-5315
Volume 7, Issue 2, 2024 (pp. 97-114)
www.abjournals.org

MACHINE LEARNING MODEL FOR ADVERSE DRUG REACTION DETECTION BASED ON NAIVE BAYES AND XGBOOST ALGORITHM

Blessing Ekong[1], Anthony Edet[2*], Uduakobong Udonna[3],

Anietie Uwah[4], and Ndueso Udoetor[5]

[1]Department of Computer Science, Akwa Ibom State University, Mkpat Enin, Nigeria.
Email: blessingekong@aksu.edu.ng

[2]Department of Computer Science, Akwa Ibom State University, Mkpat Enin, Nigeria.
Email: anthonyedet73@gmail.com

[3]Department of Computer Science, Akwa Ibom State Polytechnic, Ikot Ekpene, Nigeria.
Email: princessiswill77@gmail.com

[4]Department of Computer Science, National Open University of Nigeria, Abuja, Nigeria.
Email: uwahanietie@gmail.com

[5]Department of Computer Science, Akwa Ibom State University, Mkpat Enin, Nigeria.
Email: udoetorndueso55@gmail.com

[*]Corresponding Author's Email: anthonyedet73@gmail.com

Cite this article:

Blessing E., Anthony E., Uduakobong U., Anietie U., Ndueso U. (2024), Machine Learning Model for Adverse Drug Reaction Detection Based on Naive Bayes and XGBoost Algorithm. British Journal of Computer, Networking and Information Technology 7(2), 97-114. DOI: 10.52589/BJCNIT-35MFFBC6

Manuscript History

Received: 12 Mar 2024

Accepted: 27 May 2024

Published: 18 Jul 2024

ABSTRACT: *Adverse drug effects, commonly referred to as adverse drug reactions (ADRs), represent undesirable and unintended responses to medications or pharmaceutical products when used at recommended doses for therapeutic purposes. These effects can range from mild, tolerable symptoms to severe, life-threatening conditions and can manifest in various ways, affecting different organ systems within the human body. ADE analysis plays a pivotal role in prioritizing patient safety. By meticulously examining the relationship between drug administration and patient responses, healthcare providers can tailor medications to individual profiles, minimizing risks of adverse reactions. This ensures a patient-centric approach to treatment, where prescriptions are finely tuned to maximize efficacy while minimizing potential harm. This research aims to address this challenge by developing a machine learning system utilizing the Naive Bayes and XGBoost algorithms to enhance the categorization of drugs with adverse effects, ultimately contributing to improved patient safety and healthcare decision-making. In our approach, we made a system that detects ADR to effectively combine and collate patient medical history and drug information to detect if a patient would suffer adverse effects or reaction after taking the medication in its correct expert prescribed dose. The XGBoost algorithm gave a 75% accuracy score while Naive Bayes algorithm gave a score of 99%.*

KEYWORDS: Drug, Adverse Effect, Naive Bayes, Healthcare, Targeted Treatment.

Article DOI: 10.52589/BJCNIT-35MFFBC6
DOI URL: https://doi.org/10.52589/BJCNIT-35MFFBC6

British Journal of Computer, Networking and Information Technology
ISSN: 2689-5315
Volume 7, Issue 2, 2024 (pp. 97-114)

www.abjournals.org

INTRODUCTION

Adverse drug effects, commonly referred to as adverse drug reactions (ADRs), represent undesirable and unintended responses to medications or pharmaceutical products when used at recommended doses for therapeutic purposes (Suraj et al., 2023). These effects can range from mild, tolerable symptoms to severe, life-threatening conditions and can manifest in various ways, affecting different organ systems within the human body. Pharmacovigilance, a fundamental component of modern healthcare, is primarily concerned with monitoring and assessing the safety of pharmaceutical products throughout their lifecycle (Jiménez-Luna et al., 2021). This imperative field aims to detect, evaluate, and mitigate the risks associated with the use of drugs, ensuring that healthcare professionals and patients are well-informed about potential adverse effects. The cornerstone of effective pharmacovigilance lies in the prompt and accurate categorization of drugs based on their adverse effects, as this categorization underpins regulatory decisions, clinical practice, and patient safety (Paul et al., 2021). In recent decades, the pharmaceutical landscape has undergone a profound transformation characterized by the rapid development and introduction of novel drugs, biologics, and medical devices. While this innovation has undoubtedly brought about significant therapeutic advancements, it has also introduced a more complex and dynamic landscape of potential risks (Sapoval et al., 2022). Adverse effects associated with pharmaceuticals can vary widely in terms of severity, rarity, and presentation, making their identification and categorization a formidable challenge for healthcare stakeholders (Kim et al., 2020). Traditionally, pharmacovigilance relied heavily on manual methods for the collection and analysis of adverse event reports.

Healthcare professionals and regulatory agencies painstakingly reviewed individual case reports to assess the causality between drugs and adverse events. While this approach has been valuable in uncovering safety concerns, it has limitations in terms of scalability, speed, and consistency (You et al., 2022). The exponential growth in the volume of healthcare data, including electronic health records, social media discussions, and patient forums, further underscores the need for automated and data-driven approaches to pharmacovigilance (Golriz et al., 2021). Machine learning, a subset of artificial intelligence (Edet & Ansa, 2023), has emerged as a powerful tool in addressing the evolving challenges of pharmacovigilance. Machine learning algorithms, such as the Naive Bayes algorithm, have the capacity to process vast datasets, identify patterns, and make probabilistic predictions based on relevant features (Adam et al., 2020). When applied to the task of categorizing drugs with adverse effects, machine learning offers the potential to enhance the efficiency and accuracy of pharmacovigilance activities (Adam et al., 2020). The Naive Bayes algorithm is particularly promising for this application due to its simplicity (Ekong et al., 2023) and effectiveness in handling high-dimensional data with discrete features. It leverages the principle of conditional independence among features, which aligns well with the multifaceted nature of adverse effect data. By building a machine learning system powered by Naive Bayes, pharmacovigilance professionals can expedite the identification of potential safety concerns, prioritize regulatory actions, and ultimately contribute to the improvement of patient care.

This study endeavors to explore the utilization of the Naive Bayes algorithm within the domain of pharmacovigilance to develop a robust and efficient system for categorizing drugs based on their adverse effects (Umoren et al., 2023). By doing so, it seeks to address the critical need for enhanced methods in drug safety assessment, ultimately promoting safer and more informed pharmaceutical practices for the benefit of patients and the healthcare community at large. The

Article DOI: 10.52589/BJCNIT-35MFFBC6
DOI URL: https://doi.org/10.52589/BJCNIT-35MFFBC6

current methods of categorizing drugs based on their adverse effects in pharmacovigilance are labor-intensive, time-consuming, and often prone to inconsistency. With the rapid growth of pharmaceutical innovations and healthcare data, there is an urgent need for a more efficient and automated approach to accurately classify drugs according to their potential side effects.

This research aims to address this challenge by developing a machine learning system utilizing the Naive Bayes algorithm to enhance the categorization of drugs with adverse effects, ultimately contributing to improved patient safety and healthcare decision-making.

LITERATURE/THEORETICAL UNDERPINNING

In this section, we move into key concepts that serve to explain the research and reveal its significance. By elucidating these fundamental principles, we aim to provide a comprehensive theoretical background for understanding the study and the broader implications of its findings.

Machine Learning in Healthcare

Machine learning, a subset of artificial intelligence, has revolutionized the healthcare industry by offering innovative solutions for a wide range of challenges. It involves the development of algorithms that enable computers to learn and make predictions or decisions without being explicitly programmed. In healthcare, the application of machine learning has the potential to enhance diagnostics, treatment recommendations, patient outcomes, and overall healthcare delivery (Sorkun et al., 2020). One of the most significant advantages of machine learning in healthcare is its ability to analyze vast amounts of medical data quickly and accurately. Electronic health records (EHRs), medical imaging, genetic information, and patient-generated data provide a wealth of information that can be leveraged to improve healthcare outcomes. Machine learning algorithms can sift through these data sources to identify patterns, detect anomalies, and predict potential health issues, enabling early intervention and personalized care (Sorkun et al., 2020). Machine learning has found numerous applications in medical imaging. Deep learning algorithms, particularly convolutional neural networks (CNNs), have demonstrated remarkable success in tasks such as image classification, object detection, and segmentation. In radiology, for instance, CNNs can assist in the detection of abnormalities in X-rays, MRIs, and CT scans, aiding radiologists in making more accurate diagnoses and reducing missed diagnoses. Another critical area where machine learning shines is in disease prediction and risk stratification. By analyzing historical patient data, including demographics, medical history, and lifestyle factors, machine learning models can predict a patient's risk of developing certain diseases. This allows healthcare providers to target preventive interventions more effectively and allocate resources efficiently, ultimately improving population health (Sapoval et al., 2022).

Drug discovery and development have also benefited from machine learning. Predicting the efficacy and safety of potential drug compounds, identifying new drug candidates, and optimizing clinical trial designs are all areas where machine learning is making a significant impact. This technology accelerates the drug discovery process, potentially reducing costs and bringing life-saving medications to market more swiftly. In clinical decision support, machine learning assists healthcare professionals by providing evidence-based recommendations for diagnosis and treatment. By analyzing patient data and the latest medical literature, these

Article DOI: 10.52589/BJCNIT-35MFFBC6
DOI URL: https://doi.org/10.52589/BJCNIT-35MFFBC6

systems can suggest treatment options, identify potential drug interactions, and offer guidance in complex medical cases (Edet et al., 2024). This aids clinicians in delivering more precise and personalized care to their patients (Sapoval et al., 2022). However, the adoption of machine learning in healthcare is not without its challenges. Privacy concerns, data security, regulatory compliance, and the need for transparent and interpretable models are some of the issues that must be addressed. Additionally, there is a need for ongoing collaboration between data scientists, healthcare professionals, and policymakers to ensure that machine learning technologies are effectively integrated into clinical practice while maintaining patient safety and ethical standards (Sapoval et al., 2022). Machine learning holds immense promise in transforming the healthcare landscape. Its ability to analyze and interpret complex medical data can enhance diagnostics, treatment planning, and patient outcomes. While challenges exist, the ongoing development and integration of machine learning in healthcare have the potential to revolutionize the industry, ultimately improving the quality and efficiency of healthcare delivery for individuals worldwide (Sapoval et al., 2022). Machine learning can play a crucial role in the detection of adverse drug reactions (ADRs) by leveraging its capabilities to process and analyze vast amounts of patient data, clinical records, and drug information.

Adverse Event Prediction

Machine learning models can predict which patients are more likely to experience specific ADRs based on their medical history, genetics, and medication profiles. By assessing individual risk factors, these models can help healthcare providers make informed decisions when prescribing medications. For example, if a patient is identified as having a higher risk of a severe allergic reaction to a particular drug, the healthcare provider may choose a safer alternative or closely monitor the patient during the course of treatment (Sharma et al., 2022). Machine learning-driven ADR detection offers the potential to significantly enhance patient safety by reducing the occurrence of unexpected and severe adverse reactions. It empowers healthcare professionals to make evidence-based decisions when prescribing medications and provides a proactive approach to managing ADRs. However, it is crucial to integrate these machine learning systems into clinical workflows and ensure that they adhere to regulatory standards and ethical considerations to maintain patient privacy and safety throughout the process (Sharma et al., 2022).

The Naive Bayes Classifier

The Naive Bayes classifier is a simple yet powerful probabilistic machine learning algorithm used for classification tasks in a wide range of applications. Its simplicity and efficiency have made it a popular choice for tasks like spam email detection, sentiment analysis, and text categorization (Jiménez-Luna et al., 2021). At its core, the Naive Bayes classifier relies on Bayes' theorem (Ekong et al., 2022), which is a probabilistic formula that describes the probability of an event based on prior knowledge and evidence. The "naive" aspect of this classifier stems from its simplifying assumption that features (or variables) are conditionally independent, given the class label. This assumption, while often unrealistic in practice, allows the algorithm to work efficiently and makes it computationally tractable even for high-dimensional datasets (Jiménez-Luna et al., 2021). It requires relatively small amounts of training data and can operate well in high-dimensional feature spaces, making it suitable for tasks with many variables or features. Additionally, it is computationally efficient and can make predictions quickly, making it a valuable tool for real-time or large-scale applications.

Article DOI: 10.52589/BJCNIT-35MFFBC6
DOI URL: https://doi.org/10.52589/BJCNIT-35MFFBC6

British Journal of Computer, Networking and Information Technology
ISSN: 2689-5315
Volume 7, Issue 2, 2024 (pp. 97-114)
www.abjournals.org

Probability Theory and Bayesian Inference

Probability theory and Bayesian inference are foundational concepts in the field of statistics and have wide-ranging applications across various domains, including machine learning (Uwah & Edet, 2024), economics, physics, and healthcare. These concepts provide a systematic framework for reasoning about uncertainty, making predictions, and updating beliefs in light of new evidence. Probability theory is the mathematical study of uncertainty and randomness. At its core, it deals with the likelihood of events occurring within a given set of possibilities (Ekong et al., 2022). In probability theory, an event is typically denoted as an outcome or result of an experiment. The probability of an event is a measure of how likely it is to occur and is represented as a number between 0 and 1, where 0 indicates impossibility, 1 indicates certainty, and the values in between represent degrees of likelihood.

Existing Approach

Ahmad et al. (2022) proposed a work on Adverse Effects of COVID-19 Vaccination: Machine Learning and Statistical Approach to Identify and Classify Incidences of Morbidity and Post Vaccination Reactogenicity. The authors were motivated by the report that not all persons are eligible for COVID-19 vaccine, so they took up the research to build a machine learning model to classify people who are fit to be vaccinated based on their medical history. The authors noted that not all the persons qualified to take the vaccine because of the constituents of the vaccine which could cause a severe adverse effect or react on the patient. The authors aim to identify possible common factors in such adverse reactions to enable strategies that reduce the incidence of such reactions by using patient data to classify and characterise those at risk. The limitation in the work is that the authors used patient data only. They did not use drug or vaccine data; hence, the system is still not dependable since the problem is not a one-way thing. They examined and used patient medical histories and data documenting post vaccination effects and outcomes. In the new system, a model that put the two factors—patient medical history and drug data—into consideration is constructed.

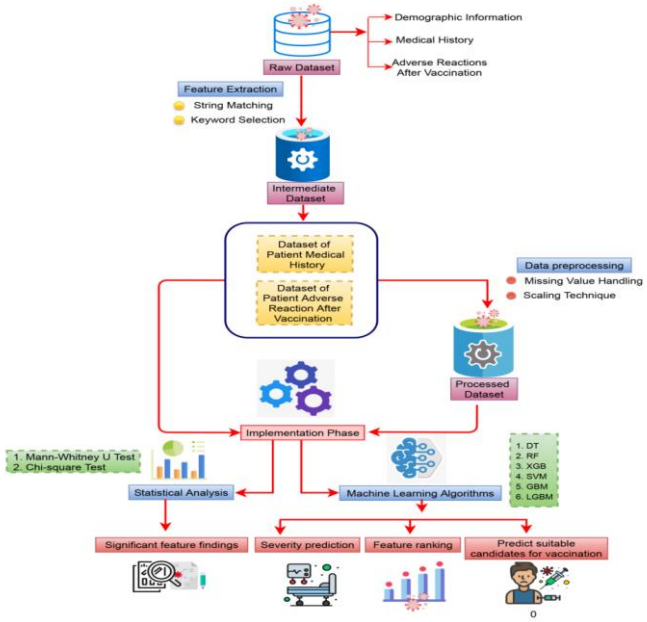

Fig. 1: Architecture of the Existing System (Ahmad et al., 2022)

Article DOI: 10.52589/BJCNIT-35MFFBC6
DOI URL: https://doi.org/10.52589/BJCNIT-35MFFBC6

British Journal of Computer, Networking and Information Technology
ISSN: 2689-5315
Volume 7, Issue 2, 2024 (pp. 97-114)
www.abjournals.org

METHODOLOGY

The current methods of categorizing drugs based on their adverse effects in pharmacovigilance are labor-intensive, time-consuming, and often prone to inconsistency. With the rapid growth of pharmaceutical innovations and healthcare data, there is an urgent need for a more efficient and automated approach to accurately classify drugs according to their potential side effects. This research aims to address this challenge by developing a machine learning system utilizing the Naive Bayes and XGBoost algorithms to enhance the categorization of drugs with adverse effects, ultimately contributing to improved patient safety and healthcare decision-making. In our approach, we made a system that detects ADR to effectively combine and collate patient medical history and drug information to detect if a patient would suffer adverse effects or reaction after taking the medication in its correct expert prescribed dose.

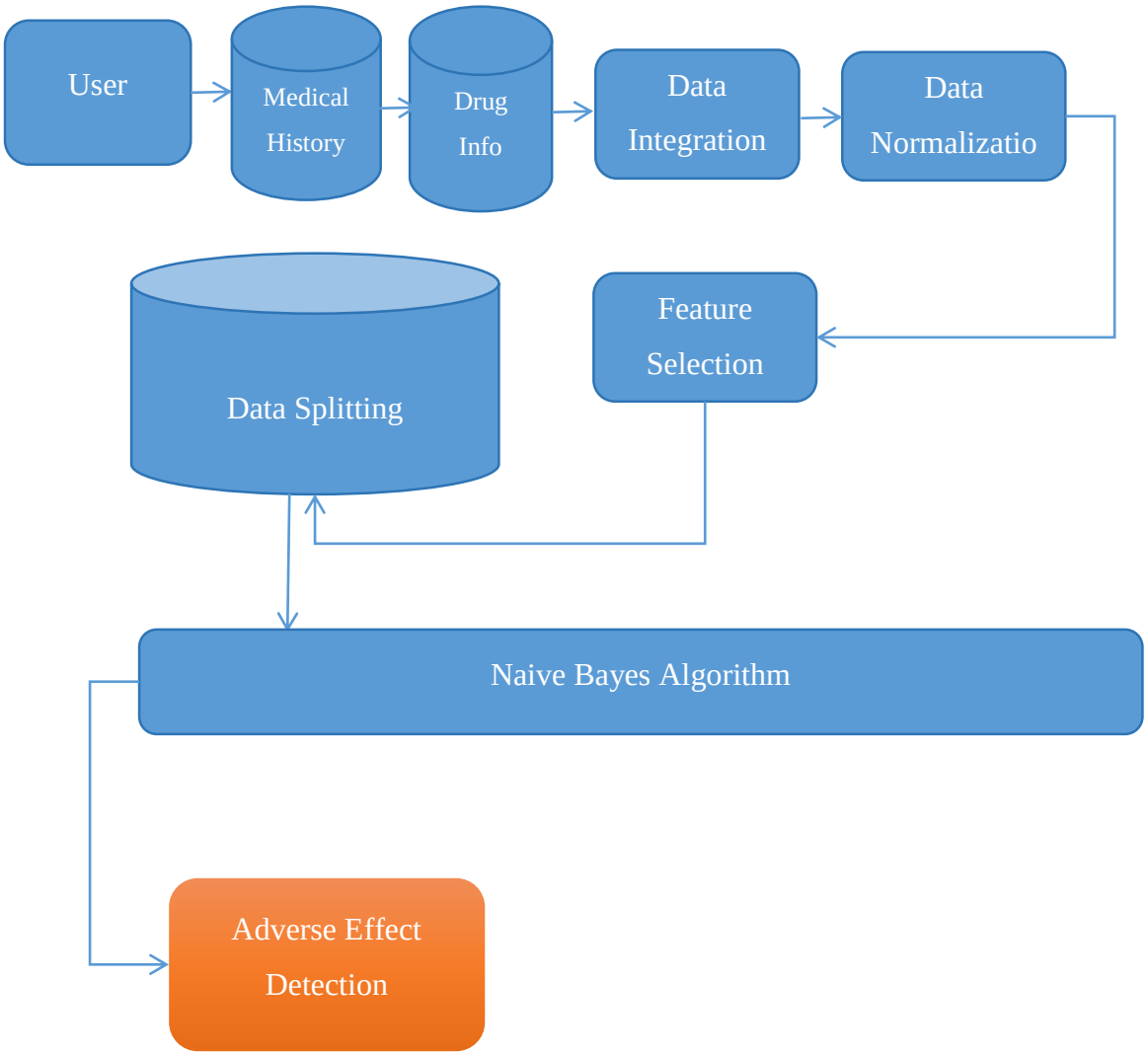

Fig 2: The Proposed Conceptual Framework

Article DOI: 10.52589/BJCNIT-35MFFBC6
DOI URL: https://doi.org/10.52589/BJCNIT-35MFFBC6

British Journal of Computer, Networking and Information Technology
ISSN: 2689-5315
Volume 7, Issue 2, 2024 (pp. 97-114)
www.abjournals.org

Data Collection

Access to high-quality datasets and resources is essential in the study of drug adverse effects, as it forms the foundation for research and analysis in this critical domain (Miljković et al., 2021). These datasets and resources enable researchers and healthcare professionals to explore and understand the adverse effects associated with various medications. One primary source of data for studying drug adverse effects is electronic health records (EHRs). EHRs contain comprehensive patient information, including prescription history, medical diagnoses, and reported adverse events. These records offer a wealth of real-world data that can be analyzed to identify patterns and associations between specific drugs and adverse reactions. EHRs have become invaluable resources for pharmacovigilance and post-marketing surveillance, allowing for the early detection of potential safety concerns associated with medications. These databases collect and store adverse event reports submitted by healthcare professionals and patients. In addition to EHRs and pharmacovigilance databases, there are publicly available datasets and resources specifically curated for research on drug adverse effects. These datasets include information on drug labels, clinical trial data, and adverse event reports, making them valuable for a wide range of studies. We leverage these resources to develop predictive models for the detection of adverse drug effects.

Data Preprocessing

The dataset was made to go through some processing activities to make it machine learning friendly. Since the dataset was gathered from different sources in order to have a huge data size as machine learning demands, data integration was done, which enabled different datasets from different sources to be put together to form one large and unique dataset. The combined dataset was then normalized to ensure usability.

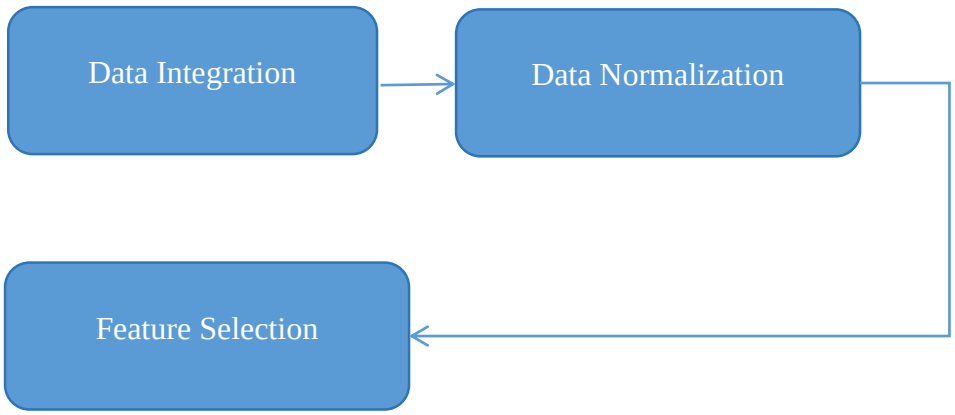

Fig. 3: Data Preprocessing Activities

Article DOI: 10.52589/BJCNIT-35MFFBC6
DOI URL: https://doi.org/10.52589/BJCNIT-35MFFBC6

British Journal of Computer, Networking and Information Technology
ISSN: 2689-5315
Volume 7, Issue 2, 2024 (pp. 97-114)
www.abjournals.org

RESULTS/FINDINGS

Table 1: XGBoost Classification Report

	Precision	Recall	F1-Score	Support
0	1.00	0.33	0.50	3
1	0.71	1.00	0.83	5
Accuracy			0.75	8
Macro avg	0.86	0.67	0.67	8
Weighted avg	0.82	0.75	0.71	8

Table 2: Naive Bayes Classification Report

	Precision	Recall	F1-Score	Support
0	0.99	0.99	0.99	29
1	0.99	0.99	0.99	47
Accuracy	0.99	0.99	0.99	76
Macro avg	0.99	0.99	0.99	76
Weighted avg	0.99	0.99	0.99	76

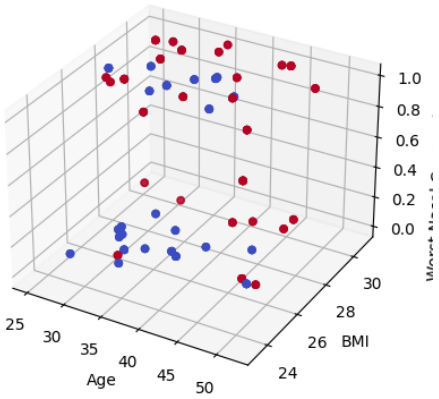

Fig. 4: Contribution of BMI, Age and Worst Nasal Congestion after Drug

The contribution of features like BMI, Age, and Worst Nasal Congestion after Drug to the classification of adverse drug effects can be understood through their impact on the predictive model. These features have a good level of influence in identifying whether a patient is likely to experience adverse effects from a drug or not. Body Mass Index (BMI) is a measure of body fat based on a person's weight and height. It is an important factor in drug metabolism and distribution in the body. Higher or lower BMI values affect how drugs are absorbed, distributed, metabolized, and eliminated (ADME).

Article DOI: 10.52589/BJCNIT-35MFFBC6
DOI URL: https://doi.org/10.52589/BJCNIT-35MFFBC6

British Journal of Computer, Networking and Information Technology
ISSN: 2689-5315
Volume 7, Issue 2, 2024 (pp. 97-114)
www.abjournals.org

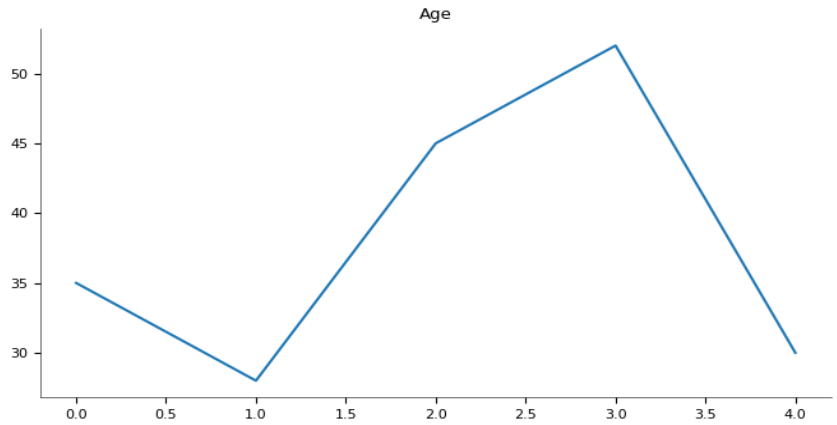

Fig. 5: Age Factor

Age is a significant factor in pharmacokinetics and pharmacodynamics—the way drugs interact with the body. The metabolism and elimination of drugs vary with age. In this research, it was discovered that children and older adults process drugs differently than young adults; these groups of people tend to experience more adverse drug effects than others.

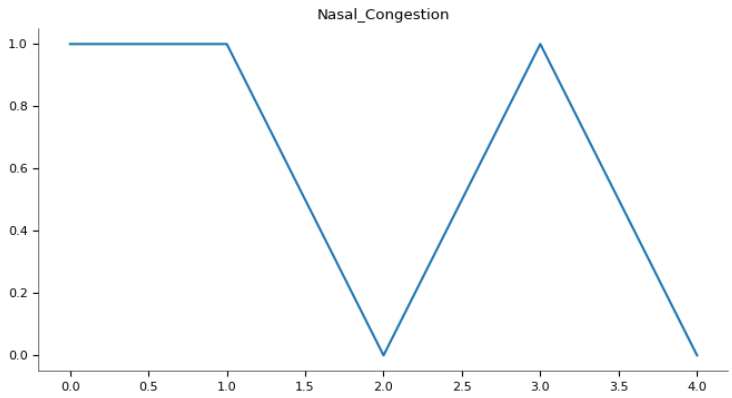

Fig. 6: Nasal Congestion Effect

This feature represents the severity of nasal congestion after taking a drug. It has been discovered that the intensity of side effects, such as nasal congestion, could be indicative of an individual's sensitivity or reaction to the drug. The model recognizes that individuals experiencing more severe nasal congestion after a drug has been administered to them are more likely to be classified as having adverse drug effects.

Article DOI: 10.52589/BJCNIT-35MFFBC6
DOI URL: https://doi.org/10.52589/BJCNIT-35MFFBC6

British Journal of Computer, Networking and Information Technology
ISSN: 2689-5315
Volume 7, Issue 2, 2024 (pp. 97-114)
www.abjournals.org

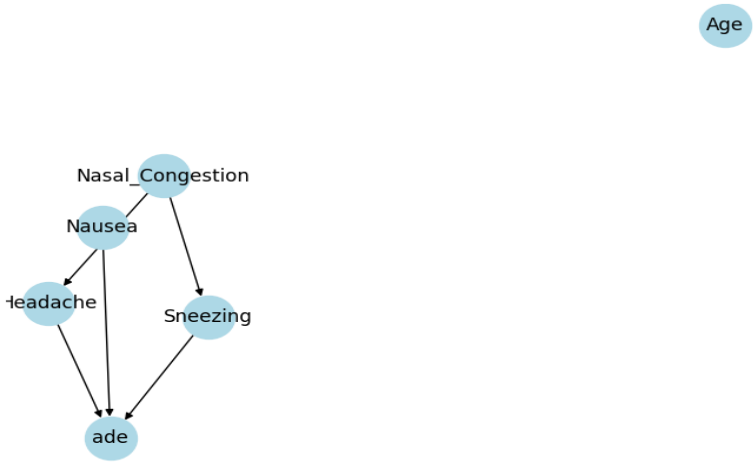

Fig. 7: Direct Acyclic Graph

A Directed Acyclic Graph (DAG) is a graph that consists of nodes connected by directed edges, where the edges have a direction and there are no cycles or loops. In other words, you cannot start at any node and follow the directed edges and eventually return to the same node.

Directed Acyclic Graph is a graph structure that provides a way to represent and understand relationships between entities, ensuring that these relationships do not form loops or cycles.

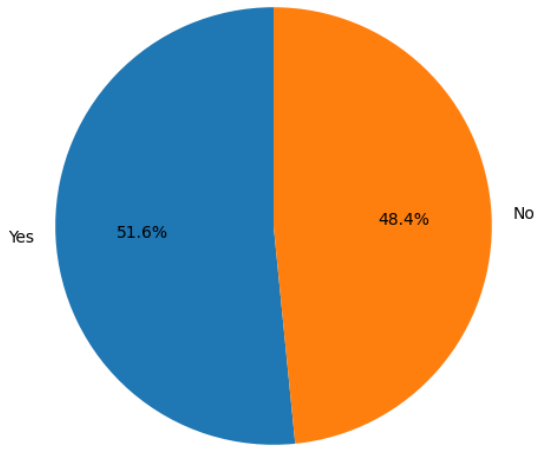

Fig. 8: Class Distribution

ade = 0: No Adverse Drug Effect

If an individual has 'ade' labeled as 0, it means that, according to the model or dataset, there is no predicted adverse drug effect for that individual. The person is not expected to experience negative reactions or side effects from the drug based on the given features or variables.

ade = 1: Adverse Drug Effect

If 'ade' is labeled as 1, it signifies that, according to the model or dataset, there is a predicted adverse drug effect for that individual. The person is expected to experience negative reactions

Article DOI: 10.52589/BJCNIT-35MFFBC6
DOI URL: https://doi.org/10.52589/BJCNIT-35MFFBC6

British Journal of Computer, Networking and Information Technology
ISSN: 2689-5315
Volume 7, Issue 2, 2024 (pp. 97-114)
www.abjournals.org

or side effects from the drug based on the provided features or variables. In simpler terms, 'ade' is a binary indicator: 0 means no predicted adverse drug effect and 1 means a predicted adverse drug effect. These predictions are made by the machine learning model based on the patterns it learned during the training process, considering various features such as age, gender, symptoms, and other relevant information in the dataset.

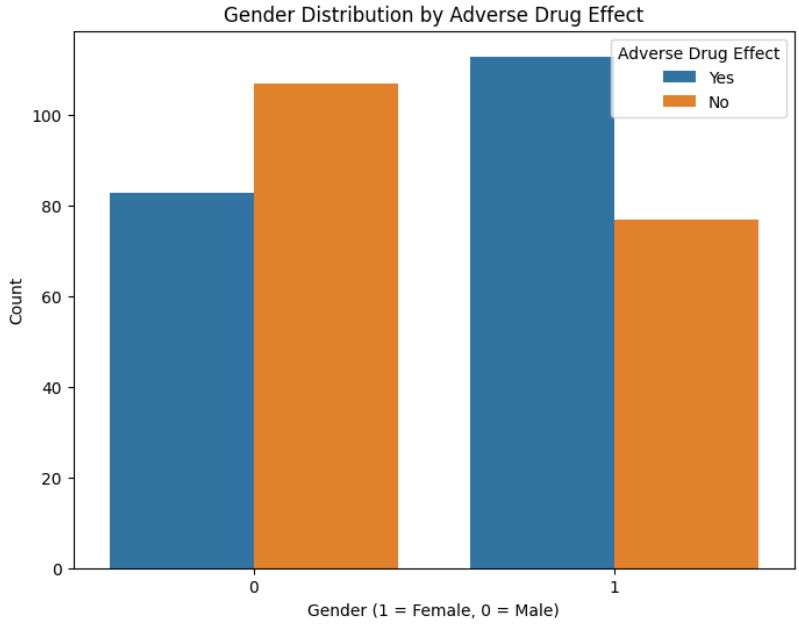

Fig. 9a: Gender Effect with Respect to Class

Figure 4.4 shows that females are more prone to adverse drug effects, that is, women react to drugs more than men.

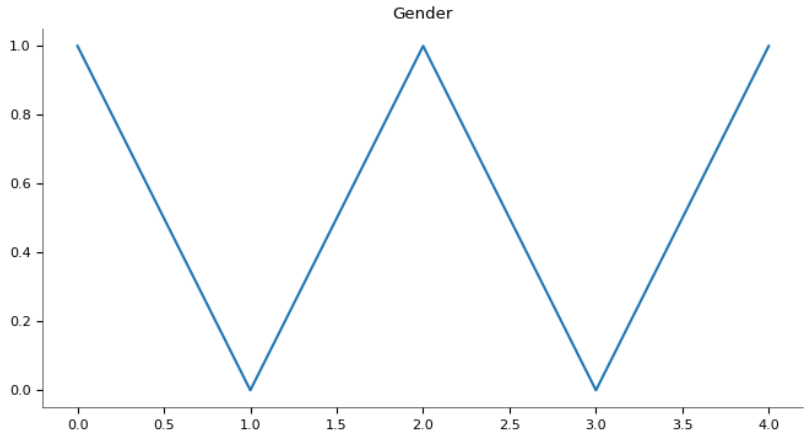

Fig. 9b: Gender Effect on Adverse Drug Effect Classification

It is advisable that extra care be taken when administering drugs to female patients.

Article DOI: 10.52589/BJCNIT-35MFFBC6
DOI URL: https://doi.org/10.52589/BJCNIT-35MFFBC6

British Journal of Computer, Networking and Information Technology
ISSN: 2689-5315
Volume 7, Issue 2, 2024 (pp. 97-114)
www.abjournals.org

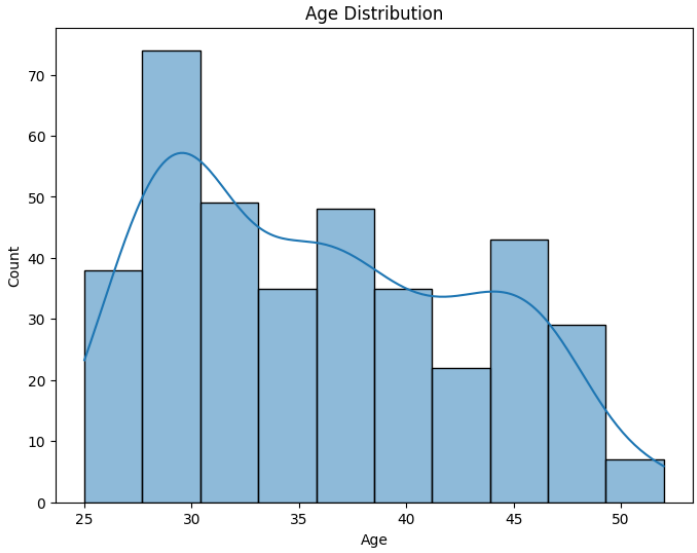

Fig. 10: Age Distribution Count

Figure 10 shows how age as a factor contributes to the problem of adverse effects of drugs.

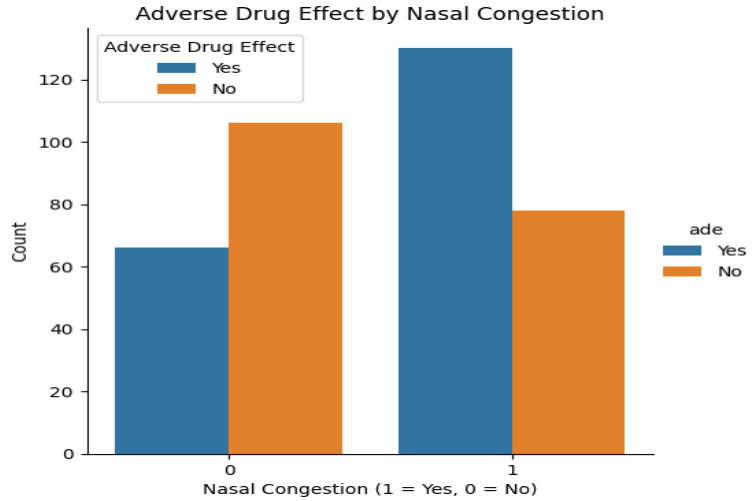

Fig. 11: Effect of Nasal Congestion

Nasal congestion is a condition characterized by a blockage or stuffiness in the nasal passages, making it difficult to breathe through the nose. It is a common symptom often associated with various respiratory infections, allergies, or irritants. Nasal congestion is a very important factor in the classification of adverse drug effects. In this work, nasal congestion after drugs shows an adverse drug effect.

Article DOI: 10.52589/BJCNIT-35MFFBC6
DOI URL: https://doi.org/10.52589/BJCNIT-35MFFBC6

British Journal of Computer, Networking and Information Technology
ISSN: 2689-5315
Volume 7, Issue 2, 2024 (pp. 97-114)

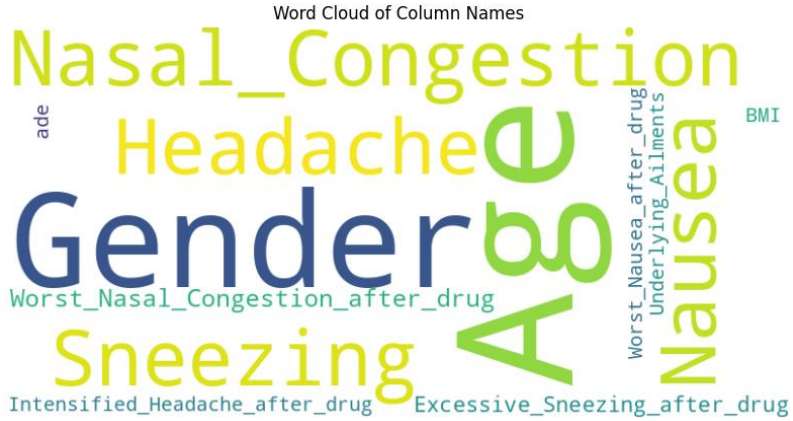

Fig. 12: Word Cloud Showing Features

Figure 12 is the word cloud of all the features in our dataset showing feature importance for the classification of adverse drug effects. A word cloud is a visual representation of text data, where the importance or frequency of each word is depicted by its size in the cloud. In a word cloud, words from a given text are arranged in a graphical manner, and the size of each word is proportional to its frequency or significance within the text.

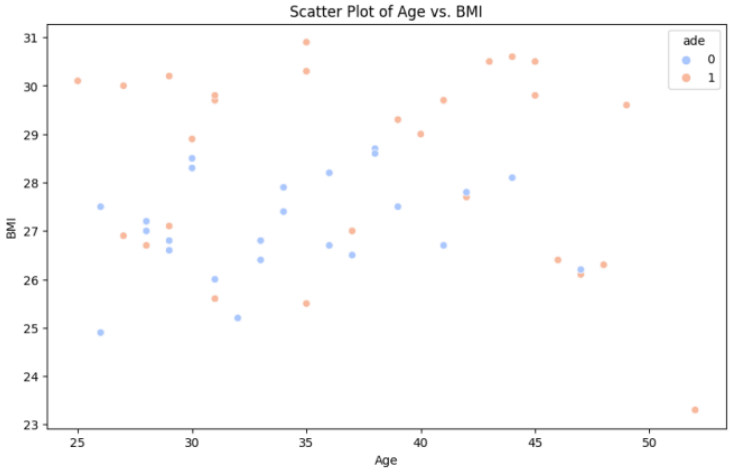

Fig. 13: BMI and Age Effect on Class

Figure 13 shows the contribution of Body Mass Index and Age on the outcome of the predicted adverse drug effect classes. Each color represents a class.

Article DOI: 10.52589/BJCNIT-35MFFBC6
DOI URL: https://doi.org/10.52589/BJCNIT-35MFFBC6

British Journal of Computer, Networking and Information Technology
ISSN: 2689-5315
Volume 7, Issue 2, 2024 (pp. 97-114)
www.abjournals.org

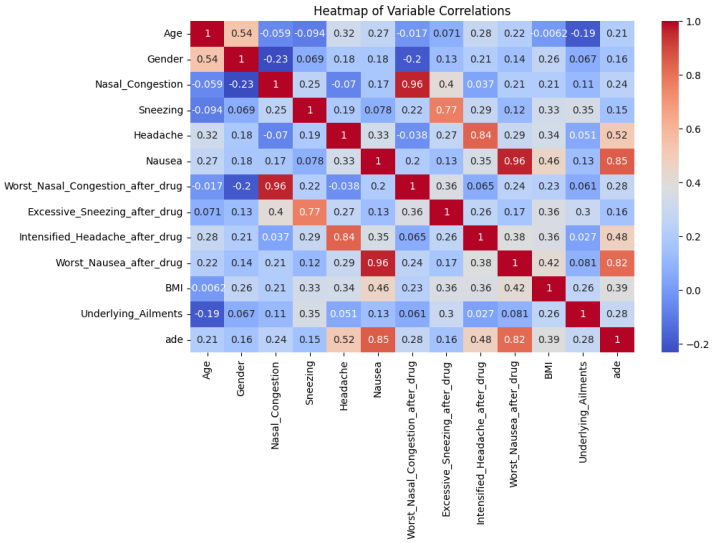

Fig. 14: Heatmap of Features Correlation

A heat map in the context of classification is a visual representation of a confusion matrix. A confusion matrix is a table that is used to evaluate the performance of a classification algorithm by summarizing the counts of true positive, true negative, false positive, and false negative predictions.

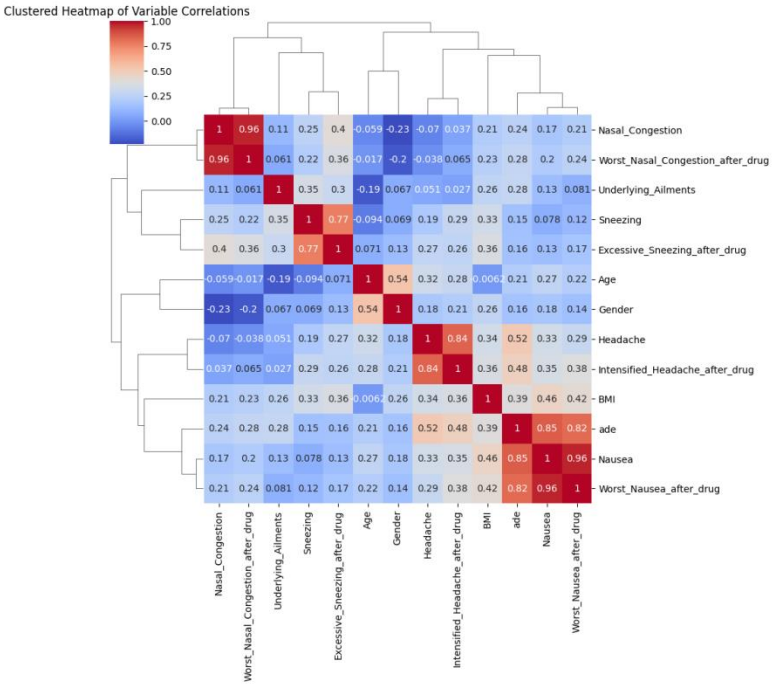

Fig. 15: Heatmap of Clustered Features Correlation

Analyzing a heat map can provide insights into the strengths and weaknesses of a classification model. It is a valuable tool for understanding the distribution of predictions and misclassifications across different classes, especially in scenarios with multiple classes.

Article DOI: 10.52589/BJCNIT-35MFFBC6
DOI URL: https://doi.org/10.52589/BJCNIT-35MFFBC6

British Journal of Computer, Networking and Information Technology
ISSN: 2689-5315
Volume 7, Issue 2, 2024 (pp. 97-114)
www.abjournals.org

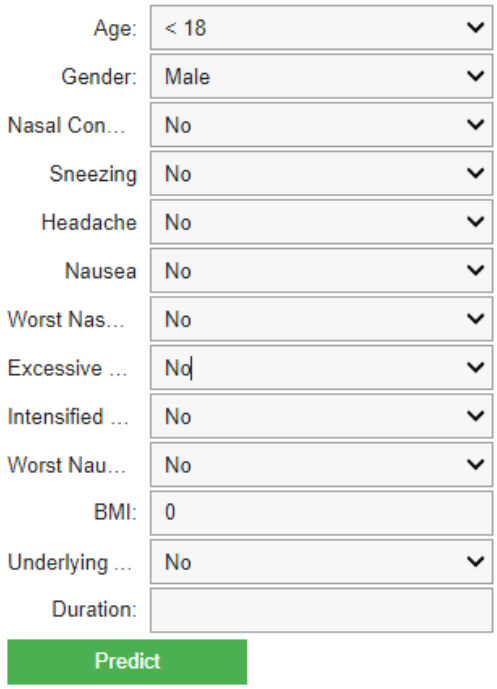

Fig 16: API for Wide Usage

DISCUSSION

The dataset contains various features that play a crucial role in assessing Adverse Drug Reaction (ADR). Among these features, factors like age and gender are fundamental determinants of how individuals may react to drugs. Age often influences the metabolism and tolerance of medications, making it an essential variable to consider. Gender differences, both biologically and in terms of hormonal variations, can impact how drugs are absorbed and processed in the body. Therefore, understanding the age and gender distribution in the dataset is crucial for a comprehensive analysis of ADR. Nasal congestion, sneezing, headache, and nausea are symptoms commonly associated with adverse drug reactions. Monitoring these symptoms is vital as they can significantly impact a patient's well-being and treatment adherence. The dataset also includes information about the intensity of these symptoms after drug administration, providing insights into the severity of the adverse effects. This data is valuable for healthcare professionals in tailoring treatment plans and minimizing the risk of adverse reactions. Worst Nasal Congestion after Drug, Excessive Sneezing after Drug, Intensified Headache after Drug, and Worst Nausea after Drug are features directly related to the severity of symptoms post-medication. Understanding these aspects is crucial for healthcare providers to assess the tolerability and safety profile of specific drugs. The dataset also includes information on Body Mass Index (BMI) and underlying ailments, factors that can influence drug metabolism and susceptibility to adverse effects. By considering these features collectively, healthcare professionals can enhance their ability to predict and manage potential adverse drug reactions, thereby improving patient safety and treatment outcomes.

Article DOI: 10.52589/BJCNIT-35MFFBC6
DOI URL: https://doi.org/10.52589/BJCNIT-35MFFBC6

British Journal of Computer, Networking and Information Technology
ISSN: 2689-5315
Volume 7, Issue 2, 2024 (pp. 97-114)

www.abjournals.org

IMPLICATION TO RESEARCH AND PRACTICE

Developing a machine learning (ML) system for detecting adverse drug reactions (ADRs) or effects holds significant importance in healthcare and pharmaceutical industries. Firstly, such systems play a pivotal role in enhancing patient safety by proactively identifying potential risks associated with various medications. By analyzing large-scale datasets comprising patient medical records, drug prescriptions, and reported adverse events, the developed model can uncover if a patient will react to a medication. Early detection of ADRs enables healthcare professionals to promptly intervene, adjust treatment plans, or even withdraw drugs to mitigate potential harm to patients. Furthermore, ML-based ADR detection systems contribute to optimizing drug development processes.

By using predictive analytics, pharmaceutical companies can gain insights into the safety profiles of candidate drugs during preclinical and clinical trials. Identifying ADRs at an early stage allows researchers to refine drug formulations, adjust dosages, or modify treatment protocols to improve drug efficacy and safety. By providing timely insights into emerging ADRs, these systems empower regulatory agencies and healthcare stakeholders to make informed decisions regarding drug approvals, labeling updates, and risk communication strategies, thereby fostering a safer and more transparent healthcare ecosystem.

CONCLUSION

ADE analysis plays a pivotal role in prioritizing patient safety. By meticulously examining the relationship between drug administration and patient responses, healthcare providers can tailor medications to individual profiles, minimizing risks of adverse reactions. This ensures a patient-centric approach to treatment, where prescriptions are finely tuned to maximize efficacy while minimizing potential harm. The result is a healthcare system that prioritizes the well-being and safety of each patient, fostering a culture of personalized medicine. The insights gained from ADE analysis contribute to the optimization of treatment plans. Recognizing patterns and correlations between patient characteristics, symptoms, and drug responses enables healthcare professionals to fine-tune treatment strategies. This individualized approach not only leads to more effective medications but also improves patient adherence. As a result, treatment outcomes are enhanced, and patients experience fewer disruptions due to adverse effects, contributing to overall healthcare efficiency. Efficient ADE analysis also aids in strategic resource allocation.

By identifying medications with lower risks for specific patient populations, unnecessary healthcare interventions can be minimized. This judicious use of resources not only benefits patients but also contributes to cost-effective healthcare. Furthermore, ADE analysis provides invaluable insights for drug developers, guiding the refinement of drug formulations and influencing the direction of future clinical trials. This dual impact on both current patient care and the future landscape of drug development underscores the significance of ADE analysis in shaping a safer and more effective healthcare ecosystem.

Article DOI: 10.52589/BJCNIT-35MFFBC6
DOI URL: https://doi.org/10.52589/BJCNIT-35MFFBC6

FUTURE RESEARCH

Future research in the area of detecting adverse drug reactions (ADRs) or effects should prioritize the development of advanced methods for integrating heterogeneous healthcare data sources and conducting comprehensive analyses. We recommend the use of data integration techniques like federated learning and data harmonization to enhance the accuracy of ADR detection algorithms.

REFERENCES

Adam G, Rampášek L, Safikhani Z, Smirnov P, Haibe-Kains B, Goldenberg A (2020) Machine learning approaches to drug response prediction: challenges and recent progress. NPJ Precis Oncol., 4:10.1038/s41698-020- 0122-1

Ahamad, M.M.; Aktar, S.; Uddin, M.J.; Rashed-Al-Mahfuz, M.; Azad, A.K.M.; Uddin, S.; Alyami, S.A.; Sarkar, I.H.; Khan, A.; Liò, P.(2023). Adverse Effects of COVID-19 Vaccination: Machine Learning and Statistical Approach to Identify and Classify Incidences of Morbidity and Postvaccination Reactogenicity. Healthcare 2023, 11, 31. https:// doi.org/10.3390/healthcare11010031.

Anietie Ekong, Blessing Ekong and Anthony Edet (2022), Supervised Machine Learning Model for EffectiveClassification of Patients with Covid-19 Symptoms Based on Bayesian Belief Network, Researchers Journal ofScience and Technology(2022),2, pp-27-33.

Edet, A. E. and Ansa, G. O. (2023). Machine learning enabled system for intelligent classification of host-based intrusion severity. Global Journal of Engineering and Technology Advances,16(03), 041–050.

Edet, A., Ekong, B. and Attih, I. (2024). Machine Learning Enabled System for Health Impact Assessment of Soft Drink Consumption Using Ensemble Learning Technique. International Journal Of Computer Science And Mathematical Theory,10(1):79-101, DOI: 10.56201/ijcsmt.v10.no1.2024.pg79.101

Ekong, B., Ekong, O., Silas, A., Edet, A., & William, B. (2023). Machine Learning Approach for Classification of Sickle Cell Anemia in Teenagers Based on Bayesian Network. Journal of Information Systems and Informatics, 5(4), 1793-1808. https://doi.org/10.51519/journalisi.v5i4.629.

Golriz Khatami S, Mubeen S, Bharadhwaj VS, Kodamullil AT, Hofmann-Apitius M, Domingo-Fernández D:(2021). Using predictive machine learning models for drug response simulation by calibrating patient-specific pathway signatures. NPJ Syst Biol Appl., 7:10.1038/s41540-021-00199-1 IEEE Trans. Ind. Informat., vol. 16, no. 9, pp. 5938–5947, Sep. 2020.jbi.2021.103838

Jiménez-Luna J, Grisoni F, Weskamp N, Schneider G(2021). Artificial intelligence in drug discovery: recent advances and future perspectives. Expert Opin Drug Discov. 16:949-59. 10.1080/17460441.2021.1909567

Kim H, Kim E, Lee I, Bae B, Park M, Nam H (2020). Artificial intelligence in drug discovery: a comprehensive review of data-driven and machine learning approaches. Biotechnol Bioprocess Eng., 25:895-930. 10.1007/s12257-020-0049-y

Article DOI: 10.52589/BJCNIT-35MFFBC6
DOI URL: https://doi.org/10.52589/BJCNIT-35MFFBC6

Miljković F, Rodríguez-Pérez R, Bajorath J (2021). Impact of artificial intelligence on compound discovery, design, and synthesis. ACS Omega. 2021, 6:33293-9. 10.1021/acsomega.1c05512

Paul D, Sanap G, Shenoy S, Kalyane D, Kalia K, Tekade RK (2021). Artificial intelligence in drug discovery and development. Drug Discov Today. 26:80-93. 10.1016/j.drudis.2020.10.010

Sapoval N, Aghazadeh A, Nute MG (2022). Current progress and open challenges for applying deep learning across the biosciences. Nat Commun., 13: 10.1038/s41467-022-29268-7

Sharma (2022). Deep Learning Model for the Automatic Classification of White Blood Cells," Computational Intelligence and Neuroscience, vol. 2022, pp. 1–13, Jan. 2022, doi: 10.1155/2022/7384131.

Sorkun MC, Astruc S, Koelman JV, Er S (2020). An artificial intelligence-aided virtual screening recipe for twodimensional materials discovery. Npj Comput Mater., 24:10.1038/s41524-020-00375-7

Suraj Tanaji Khot, Manasi Ishwar Karande, Madhubala Janardan Kamble, Manisha M Murgude(2023). Pharmacovigilance Basics Overview and Application of Artificial Intelligence in Pharmacovigilance. International journal of pharmacy and pharmaceutical research,27,(1),pp.1-15.

Umoren, I., Abe, O., Ansa, G., Inyang, S., & Umoh, I. (2023). A New Index for Intelligent Classification of Early Syndromic of Cardiovascular (CVD) Diseases Based on Electrocardiogram (ECG). *European Journal of Computer Science and Information Technology*, *11*(4), 1-21.

Uwah, A. and Edet, A. (2024).Customized Web Application for Addressing Language ModelMisalignment through Reinforcement Learning from HumanFeedback. World Journal of Innovation And Modern Technology,8,(1), 62-71. DOI: 10.56201/wjimt.v8.no1.2024.pg62.71.

You Y, Lai X, Pan Y(2022). Artificial intelligence in cancer target identification and drug discovery . Signal Transduct Target Ther., 7:10.1038/s41392-022-00994-0

Article DOI: 10.52589/BJCNIT-35MFFBC6
DOI URL: https://doi.org/10.52589/BJCNIT-35MFFBC6