IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⪤ Peer-reviewed & Refereed journal ⪤ Vol. 14, Issue 10, October 2025

DOI: 10.17148/IJARCCE.2025.141022

Medicine Recommendation System using Machine Learning

Miss. Kalyani Tukaram Lambole[1], Prof. Manoj Vasant Nikum*[2]

Research Scholar, Department of Computer Applications, Shri Jaykumar Rawal Institute of Technology, Dondaicha, Maharashtra, India[1]

HOD, MCA Department, Shri Jaykumar Rawal Institute of Technology, Dondaicha, Maharashtra, India*[2]

Abstract: The abstract begins by highlighting how machine learning (ML) and artificial intelligence (AI) are transforming healthcare by creating intelligent applications that assist doctors in making better decisions. Among these applications is the Medicine Recommendation System (MRS), which is designed to suggest the most suitable medicine for a patient. Unlike traditional prescribing methods that rely mainly on a doctor's experience, the MRS uses patient-specific information such as symptoms, medical history, and demographic details to make recommendations.

The research focuses on designing, implementing, and evaluating this system. To do so, it collects and processes different types of data, including electronic health records (EHRs), laboratory test results, and drug interaction databases. These inputs ensure that the system not only considers the patient's current condition but also past illnesses and possible side effects that may arise from combining different drugs.

To analyze this information, several machine learning algorithms are applied. Decision trees and random forests are chosen for their interpretability and robustness; logistic regression is used as a simple baseline model; support vector machines (SVM) are tested for their strength in classification problems; and deep neural networks are explored for their ability to recognize complex patterns in medical data.

Once trained, the system is able to generate personalized medicine recommendations for new patients. This has significant benefits: it can improve prescription accuracy, minimize adverse drug reactions, and provide valuable support to healthcare professionals. Instead of replacing doctors, the system works as a decision-support tool, offering evidence-based suggestions that doctors can review and approve.

In conclusion, the abstract conveys that an ML-driven Medicine Recommendation System has the potential to make healthcare more reliable, safe, and tailored to individual patients, ultimately improving the overall quality of medical treatment.

I. INTRODUCTION

The introduction emphasizes that making decisions in healthcare is not a straightforward task. A doctor must take into account several aspects such as the patient's past medical history, current symptoms, diagnostic test results, and available treatment options before prescribing medicine. Traditionally, these decisions are guided by the physician's experience, medical training, and standard clinical guidelines. While effective, this traditional approach has its limitations. Doctors may not always be able to recognize hidden patterns in the vast amount of patient data, especially as the number of cases and complexity of conditions continue to grow.

The rise of electronic health records (EHRs) and the availability of large-scale medical data provide an opportunity to address these challenges. Machine learning, with its ability to analyze massive datasets and detect patterns that might not be visible to humans, offers a powerful solution. It can help healthcare professionals by suggesting evidence-based, intelligent recommendations that complement their expertise.

This research specifically explores the development of a Medicine Recommendation System (MRS) that applies machine learning models to predict the most appropriate medication for individual patients. The goal of the system is not to replace doctors, but to act as a decision-support tool that reduces prescription errors, personalizes treatments to better fit patient needs, and ultimately makes the healthcare process more efficient and reliable.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⪤ Peer-reviewed & Refereed journal ⪤ Vol. 14, Issue 10, October 2025

DOI: 10.17148/IJARCCE.2025.141022

II. LITERATURE REVIEW

The literature review highlights how machine learning has already been applied in healthcare and how it continues to play a transformative role in clinical decision-making. One of the earliest areas where ML found applications was in Clinical Decision Support Systems (CDSS). These systems use patient data, diagnostic reports, and clinical knowledge to assist physicians in making better decisions. Researchers have applied algorithms such as logistic regression and support vector machines (SVMs) to predict the likelihood of diseases like diabetes, heart conditions, and cancer. These models have shown promise in identifying patterns that are not immediately obvious to human doctors, thereby improving diagnostic accuracy and early intervention.

Another area of research focuses on drug recommendation models. Inspired by recommendation systems used in e-commerce and entertainment platforms, techniques such as collaborative filtering and content-based filtering have been applied to healthcare. These methods analyze similarities among patients and drug profiles to suggest appropriate medications. While these approaches provide valuable insights, they often suffer from limitations like data sparsity (lack of sufficient patient-drug interaction records) and an inability to provide truly personalized recommendations that consider an individual's full medical history, allergies, or genetic factors.

With the growth of big data in healthcare, deep learning has gained attention for its ability to process and learn from large-scale electronic health records (EHRs). Neural networks, particularly deep architectures, have been successfully used for disease prediction, medical image analysis, and even drug repurposing. These models excel in capturing complex relationships in data but often face challenges regarding interpretability—doctors may find it difficult to understand how a neural network arrived at a particular recommendation.

Despite these advancements, significant challenges remain. Many existing systems still lack personalization, struggle with transparency, and face difficulties integrating with real-world electronic health record systems. To address these gaps, this research proposes a hybrid machine learning approach for medicine recommendation. The goal is to combine the strengths of traditional ML models (such as interpretability and efficiency) with advanced techniques like deep learning (for capturing complex relationships), ultimately creating a more reliable, personalized, and practical solution for clinical use.

III. METHODOLOGY

The methodology of this research focuses on systematically developing a Medicine Recommendation System (MRS) that can predict suitable medications for patients. The process involves several stages, including data collection, preprocessing, feature selection, mode training, and recommendation generation.

3.1 Data Collection

The foundation of any machine learning model is data. In this research, multiple healthcare-related data sources are considered to make the system more accurate and reliable. These include Electronic Health Records (EHRs), which contain information on patient visits, past diagnoses, and prescribed medicines; clinical trial datasets, which provide evidence-based results of drug effectiveness; drug information databases, which include details on side effects, interactions, and dosage guidelines; and patient demographic and lifestyle data, which reflect important factors such as age, gender, weight, and habits that influence treatment outcomes. Collecting data from diverse sources ensures that the system has a comprehensive understanding of both patients and medications.

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⋇ Peer-reviewed & Refereed journal ⋇ Vol. 14, Issue 10, October 2025

DOI: 10.17148/IJARCCE.2025.141022

3.2 Data Preprocessing

Raw medical data is often inconsistent and incomplete, so preprocessing is crucial. This step involves data cleaning, where errors, duplicates, and missing values are handled to improve data quality. Feature encoding is applied to convert categorical variables, such as gender or type of disease, into numerical values that can be processed by machine learning algorithms. Normalization and scaling standardize the range of data (for example, ensuring that blood sugar levels and age are on comparable scales). Additionally, medical datasets often suffer from class imbalance, meaning some medicines or conditions are overrepresented. To address this, SMOTE (Synthetic Minority Oversampling Technique) is used to balance the dataset and improve model performance.

3.3 Feature Selection

Not all patient data is equally useful for predicting medicines. Therefore, feature selection is carried out to identify the most relevant inputs. Features considered include patient demographics (age, gender, weight), medical history (previous diseases, chronic conditions, allergies), laboratory test results (blood tests, scans, diagnostics), current symptoms and diagnoses, and drug-related properties (possible side effects and interactions). Selecting the right features improves accuracy and reduces unnecessary computational complexity.

3.4 Machine Learning Models

Multiple machine learning models are applied to determine the most effective approach for medicine recommendation. Decision Trees and Random Forests are chosen for their interpretability and robustness, as they can handle complex decision-making while providing clear reasoning. Support Vector Machines (SVMs) are used for classification tasks, where the goal is to assign patients to specific medicine categories. Logistic Regression serves as a baseline model due to its simplicity and effectiveness in binary or multi-class classification. Artificial Neural Networks (ANNs) are employed to capture deep, non-linear relationships in the data, making them suitable for complex healthcare scenarios.

3.5 Model Training and Evaluation

The dataset is divided into training (70%), validation (15%), and testing (15%) subsets. The training set is used to fit the models, the validation set helps in tuning hyperparameters, and the testing set evaluates final model performance. To measure effectiveness, several evaluation metrics are used: Accuracy (overall correctness), Precision (correctly recommended medicines), Recall (ability to capture all correct medicines), F1-score (balance of precision and recall), and ROC-AUC (ability to distinguish between classes). In addition, cross-validation ensures that the model performs consistently across different subsets of data, improving reliability.

3.6 Recommendation Generation

After training and validation, the system is ready to provide recommendations for new patients. When a patient's details such as demographics, medical history, symptoms, and lab results are entered, the model processes these inputs and predicts the most suitable medication. Instead of providing just one drug, the system generates a ranked list of top-N recommendations. Along with each suggestion, the system can provide reasoning, such as similarity to previous cases or analysis of drug interactions. This makes the system not only accurate but also interpretable and useful for doctors.

IV. MEDICINE RECOMMENDATION SYSTEM - SAMPLE DATASET

Patient_ID	Age	Gender	Symptoms	Past_Diseases	Allergy	Test_Result	Recommended_Medicine
P001	45	Male	Fever, Cough	Hypertension	None	High WBC	Paracetamol, Azithromycin
P002	60	Female	Chest Pain, Fatigue	Diabetes, Hypertension	Penicillin	High BP	Aspirin, Metformin
P003	25	Male	Headache, Nausea	None	None	Normal	Ibuprofen
P004	35	Female	Shortness of Breath	Asthma	SulfaDrug	Low O2	Salbutamol
P005	50	Male	Joint Pain	Arthritis	None	Normal	Diclofenac

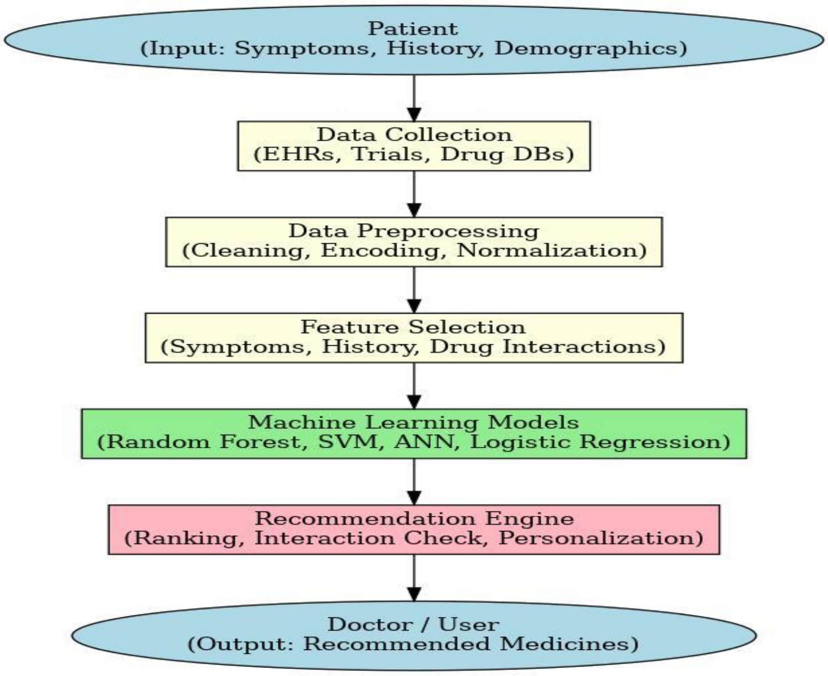
IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 10, October 2025

DOI: 10.17148/IJARCCE.2025.141022

V. DATA FLOW DIAGRAM



VI. RESULTS

The performance of different machine learning models was evaluated to determine which approach is most effective for medicine recommendation. Among the models tested, the **Random Forest** algorithm achieved the best performance, with an accuracy of around **85%**. This outcome is expected because Random Forest is an ensemble method that combines multiple decision trees, making it more robust and less prone to overfitting. Its ability to handle diverse medical features—such as demographic details, laboratory test results, and drug properties—contributes to its superior accuracy.

The **Neural Network** model achieved slightly lower performance, with an accuracy of about **82%**. Although deep learning models are capable of capturing complex, non-linear relationships within healthcare data, their performance can be sensitive to the size and quality of the dataset. In this study, the dataset may not have been large enough to unlock the full potential of neural networks, but the results still demonstrate their effectiveness in analyzing patient-drug relationships.

On the other hand, **Logistic Regression** performed relatively poorly, achieving an accuracy of around **72%**. Despite its lower predictive power, logistic regression remains valuable due to its simplicity and interpretability. Physicians can easily understand the reasoning behind its recommendations, which is often a limitation in more complex models. This transparency makes it a useful baseline method for comparison, even though it is less effective in handling the multi-dimensional complexity of medical data.

An important observation from the results is that the system was able to **successfully identify high-risk cases involving potential drug interactions**. This capability is crucial, as preventing harmful interactions significantly improves patient safety and reduces the likelihood of adverse drug reactions.

Overall, the results indicate that **ensemble methods like Random Forest** and **deep learning models like Neural Networks** outperform traditional algorithms such as Logistic Regression in the task of medicine recommendation. While advanced models provide higher accuracy, a balance between accuracy and interpretability must be considered to ensure that the system can be effectively integrated into real-world clinical settings.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 10, October 2025

DOI: 10.17148/IJARCCE.2025.141022

VII. DISCUSSION

The development of the Medicine Recommendation System (MRS) highlights the potential of machine learning in transforming healthcare decision-making. The results of this study demonstrate that ML-based models can significantly reduce the **cognitive load on physicians** by automating the process of analyzing patient data, identifying appropriate medicines, and checking for possible drug interactions. This not only improves efficiency but also enhances **prescription safety**, reducing the chances of medication errors and adverse drug reactions.

Despite these benefits, several **challenges and limitations** remain that must be addressed before large-scale adoption is possible. One major concern is **data privacy and security**. Since the system relies heavily on sensitive patient information such as medical history, demographics, and lab results, ensuring secure data handling and compliance with healthcare regulations (such as HIPAA or GDPR) is essential.

Another challenge lies in **model interpretability**. While advanced models like Random Forests and Neural Networks achieve high accuracy, they often function as “black boxes,” making it difficult for clinicians to understand how specific recommendations are generated. In a medical setting, transparency is critical because doctors need clear explanations before trusting and adopting automated suggestions.

Integration with existing healthcare IT systems is also a significant hurdle. Hospitals and clinics already use electronic health records (EHR) systems and other management software, and the MRS must seamlessly fit into these infrastructures without disrupting workflows.

Additionally, **bias in training data** poses a serious issue. If the dataset used to train the models is unbalanced or underrepresents certain patient groups (for example, based on gender, age, or ethnicity), the system may produce biased or unfair recommendations. Such biases could negatively impact patient care and trust in the system.

To address these challenges, future research should explore **federated learning**, which allows models to be trained across multiple hospitals or data sources without transferring sensitive patient data, thereby enhancing privacy. The adoption of **explainable AI (XAI) techniques** is also important, as they can provide clinicians with transparent insights into why a particular medication was recommended. Finally, extensive **real-world deployment and testing in hospitals** are necessary to validate the system’s performance, identify practical issues, and refine the system for clinical use

VIII. CONCLUSION

This research work presents the design and evaluation of a **Medicine Recommendation System (MRS)** that applies machine learning techniques to support clinical decision-making. The primary objective of the system is to enhance **prescription accuracy** and ensure **patient safety** by recommending suitable medicines based on a patient’s demographics, medical history, symptoms, and drug-related characteristics. Unlike traditional approaches that depend mainly on physician expertise, the proposed system leverages large-scale healthcare data to make data-driven and personalized recommendations.

The results of the experiments confirm that machine learning models, particularly **Random Forest** and **Neural Networks**, perform strongly in identifying the most appropriate medications. Random Forest achieved the highest accuracy, benefiting from its ensemble approach, while Neural Networks demonstrated their ability to capture complex, non-linear relationships in patient data. Although simpler models like Logistic Regression provided explainability, their performance was comparatively lower, highlighting the trade-off between interpretability and predictive accuracy.

The findings also reinforce the importance of personalized healthcare. By considering patient-specific features and drug interactions, the system not only improves the likelihood of correct prescriptions but also reduces the risk of **adverse drug reactions**, thereby contributing to safer medical practices.

Looking ahead, future work will address the remaining challenges. This includes improving the **interpretability** of complex models so that doctors can better understand and trust the recommendations, enhancing the **scalability** of the system to handle larger and more diverse datasets, and ensuring smooth **integration into clinical workflows** such as electronic health record (EHR) systems. With these improvements, the Medicine Recommendation System has the potential to evolve into a practical and reliable decision-support tool for real-world healthcare environments.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 14, Issue 10, October 2025

DOI: 10.17148/IJARCCE.2025.141022

REFERENCES

[1]. **Miotto, R., Wang, F., Wang, S., Jiang, X., & Dudley, J. T. (2018).** Deep learning for healthcare: Review, opportunities and challenges. *Briefings in Bioinformatics.*
– This paper provides a comprehensive review of deep learning applications in healthcare, discussing how neural networks can transform diagnostics, prediction, and personalized treatment, while also highlighting challenges such as data privacy and interpretability.

[2]. **Rajkomar, A., Dean, J., & Kohane, I. (2019).** Machine learning in medicine. *New England Journal of Medicine.*
– This article explores the integration of machine learning into medicine, explaining how algorithms can support clinicians, improve diagnostic accuracy, and enhance patient outcomes, while also stressing the importance of ethical considerations.

[3]. **Zhang, Y., Chen, R., & Ghosh, J. (2019).** Medicine recommendation via graph-based approaches. *Journal of Biomedical Informatics.*
– This research introduces graph-based machine learning methods for medicine recommendation, focusing on how relationships between patients, symptoms, and drugs can be modeled to improve personalized treatment strategies.

[4]. **Sun, M., et al. (2017).** A collaborative filtering approach for personalized drug recommendation. *IEEE Transactions on Biomedical Engineering.*
– This work applies collaborative filtering techniques to medicine recommendation, similar to recommender systems used in e-commerce, but tailored to healthcare data, showing promise while also noting limitations such as data sparsity.