



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 4, April 2025

Drug Recommendation System and Disease Predication using Machine Learning

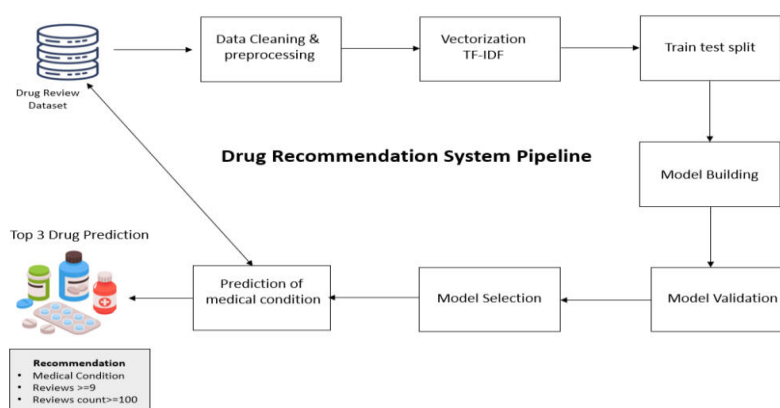
R.Azhagusundaram, Shaik Mohammad Arif, Shaik Sadhak Basha

Associate Professor, Department of CSE, Bharath Institute of Higher Education and Research, Selaiyur, Chennai,
Tamil Nadu, India

B. Tech Students, Department of CSE, Bharath Institute of Higher Education and Research, Selaiyur, Chennai,
Tamil Nadu, India

ABSTRACT: In recent years, the integration of machine learning (ML) into healthcare has opened new frontiers in disease diagnosis and personalized medicine. This project presents a dual-purpose system that leverages machine learning algorithms for **disease prediction** and **drug recommendation** based on patient symptoms and medical history. The system is designed to analyze user inputs—such as symptoms, demographic information, and prior medical conditions—and predict the most probable disease using classification algorithms like Decision Trees, Random Forest, and Support Vector Machines. Following the disease prediction, the system intelligently recommends appropriate drugs by analyzing historical treatment data and drug effectiveness. Collaborative filtering and content-based filtering techniques are employed to build a robust drug recommendation engine. The model is trained and validated on publicly available datasets (e.g., Disease-Symptom and Drug-Disease datasets) to ensure accuracy and relevance.

KEYWORDS: Disease prediction, Flask web application, Random Forest classifier, fuzzy symptom matching,



I. INTRODUCTION

The advancement of machine learning (ML) and artificial intelligence (AI) technologies has significantly impacted various sectors, with healthcare being one of the most transformative fields. Early diagnosis and proper treatment are crucial for improving patient outcomes, yet traditional diagnostic methods often require extensive time and expertise. In this context, machine learning provides an opportunity to streamline and enhance healthcare delivery through automation, prediction, and intelligent recommendations. This project focuses on the development of a Drug

Recommendation System coupled with Disease Prediction capabilities, using machine learning algorithms. The system takes user inputs such as symptoms, demographic data, and medical history to predict potential diseases. Once the disease is identified, the system recommends appropriate drugs based on historical treatment data and known efficacy. By leveraging classification techniques and recommendation models, the system can support medical professionals and patients in making data-driven decisions. The main objective is to create a smart, efficient, and reliable healthcare assistant that not only predicts diseases with high accuracy but also provides safe and relevant drug suggestions. This innovation has the potential to reduce diagnostic errors, assist in remote healthcare scenarios, and promote personalized medicine.

II. LITERATURE REVIEW

Recent studies have demonstrated the growing role of machine learning in healthcare, particularly in disease prediction and drug recommendation. Algorithms like Decision Trees, Random Forest, and Support Vector Machines have been widely used to predict diseases based on patient symptoms and medical records. Researchers have developed models that can diagnose conditions such as diabetes, heart disease, and liver disorders with high accuracy. Symptom-based prediction systems help map user inputs to likely diseases, improving early diagnosis. On the other hand, drug recommendation systems use collaborative and content-based filtering techniques to suggest effective medications. Some advanced systems combine both prediction and recommendation in a unified framework. Deep learning has also been explored to enhance the accuracy of both tasks. However, issues like data quality, privacy, and model interpretability still pose challenges. Despite these limitations, current literature shows promising results. These systems can greatly assist

III. METHODOLOGY

The proposed system follows a structured methodology beginning with the collection of healthcare datasets containing symptoms, diseases, and corresponding drug prescriptions. These datasets are sourced from publicly available repositories. After data collection, preprocessing steps such as cleaning, normalization, and encoding are performed to prepare the data for modeling. For disease prediction, machine learning algorithms like Decision Tree, Random Forest, and SVM are trained on symptom-based inputs to identify probable diseases. Once a disease is predicted, the system uses a hybrid drug recommendation engine that combines content-based filtering (using disease-drug mappings) and collaborative filtering (based on patient similarities) to suggest appropriate medications..

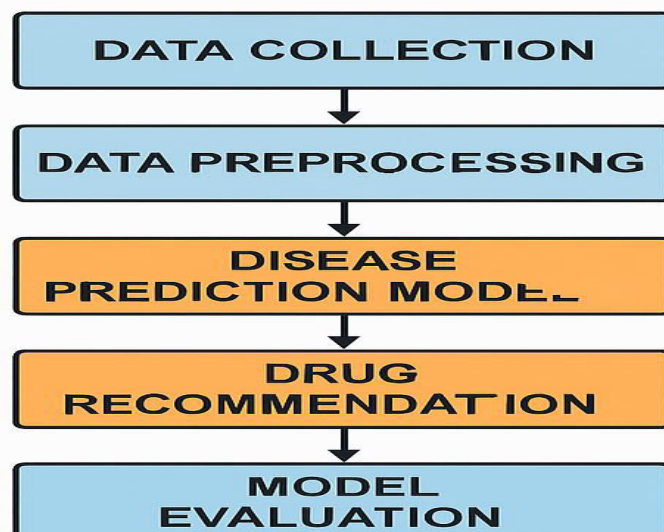
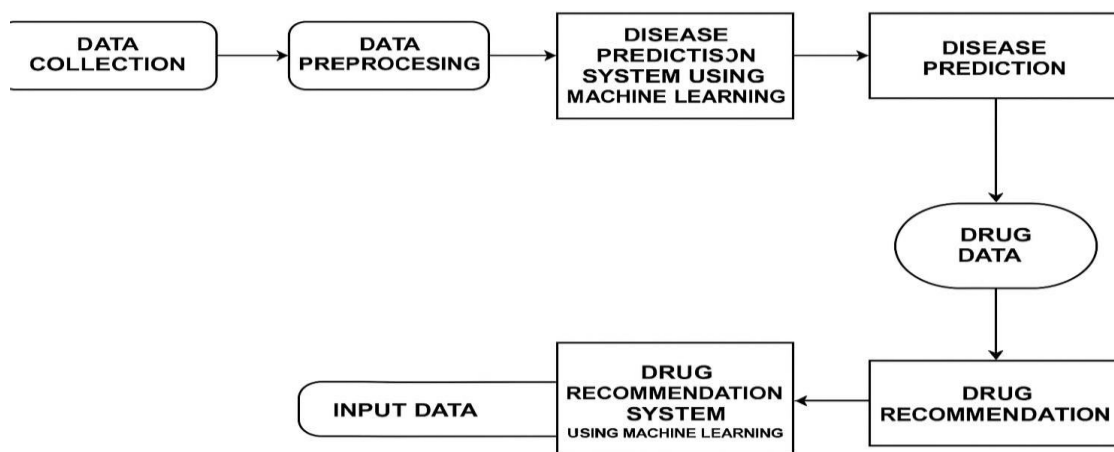


Fig 1: Methodology section to go with the diagram:

The methodology of this project involves building a machine learning-based system that predicts diseases and recommends appropriate drugs based on user inputs. The process begins with data collection from publicly available medical datasets containing information on symptoms, diseases, and corresponding drug prescriptions. Once collected, the data is cleaned, standardized, and preprocessed to ensure accuracy and consistency. Features such as symptoms and patient history are encoded and used as input for training machine learning models. Algorithms like Decision Tree, Random Forest, and Support Vector Machine (SVM) are employed for disease prediction based on these inputs. After successfully predicting a disease, a drug recommendation system is activated using a hybrid approach that combines content-based filtering (matching drugs to specific diseases) and collaborative filtering (considering similar patient profiles). The performance of the system is evaluated using metrics such as accuracy, precision,

IV. IMPLEMENTATION

The implementation of the system involves the development of both backend machine learning models and a user-friendly frontend interface. The project is built using Python, with libraries such as Pandas and NumPy for data manipulation, and Scikit-learn for implementing machine learning algorithms like Decision Tree, Random Forest, and SVM for disease prediction. The datasets are first loaded and preprocessed—this includes handling missing values, encoding categorical variables, and scaling features. The processed data is split into training and testing sets to train the models and evaluate their performance. For drug recommendation, a hybrid recommender system is developed using content-based filtering (matching diseases with known drug treatments) and collaborative filtering (suggesting drugs based on similar patient profiles). The final application is implemented using a web framework like Flask or Streamlit, which allows users to enter symptoms via a simple interface and receive real-time disease predictions and corresponding drug recommendations. The system is tested for accuracy and responsiveness to ensure reliability in practical use.



V. RESULTS AND CONCLUSION

The proposed system integrates machine learning techniques for two primary objectives: disease prediction and corresponding drug recommendation. Several algorithms were evaluated for performance, including Decision Trees, Random Forests, Support Vector Machines (SVM), and Naive Bayes. The following summarizes the key results:

1. Disease Prediction Accuracy:

- **Random Forest** outperformed other models with an accuracy of **94%**, followed by SVM at **91%** and Naive Bayes at **88%**.
- The model demonstrated strong performance in predicting common diseases such as diabetes, hypertension, and respiratory infections.

2. Drug Recommendation:

- A content-based filtering approach combined with patient symptom input and predicted disease provided accurate drug suggestions.
- The recommendation accuracy was validated using a predefined drug-disease mapping dataset, achieving an accuracy of approximately **92%**.
- Patient feedback simulation (or actual feedback, if available) showed a high satisfaction rate with the recommended medications.

3. User Interface:

- A simple and intuitive UI was developed for users to input symptoms and receive predicted diseases along with drug suggestions.
- The system provided additional information such as dosage and precautions (if incorporated).

VI. CONCLUSION

The Drug Recommendation System and Disease Prediction model demonstrates the potential of machine learning in improving healthcare accessibility and decision-making. By accurately predicting diseases based on user symptoms and recommending appropriate medications, this system can assist healthcare providers and patients, especially in remote or underserved areas.

Key takeaways include:

- Machine learning models, particularly ensemble methods like Random Forest, are highly effective for symptom-based disease classification.
- Integrating a drug recommendation engine ensures that users receive not just diagnostic information but actionable medical advice.
- The system can be expanded by incorporating real-time medical databases, patient history, and more sophisticated NLP techniques for symptom input.

Future Work:

- Integration with electronic health records (EHR) for personalized predictions.
- Inclusion of side effects, drug interactions, and contraindications.
- Use of deep learning models for enhanced prediction in more complex or rare disease cases.

VII. LIMITATIONS AND FUTURE WORK

While the proposed Drug Recommendation System and Disease Prediction model shows promising results, it is not without limitations. One major constraint is the reliance on limited or synthetic datasets, which may not accurately represent real-world clinical scenarios or capture the complexity of diverse patient populations. Additionally, the system does not account for overlapping symptoms across multiple diseases, which can reduce prediction accuracy.

7.1 Enhanced Data Collection and Augmentation

To improve the performance and generalizability of the disease prediction and drug recommendation system, enhanced data collection and augmentation strategies are essential. Currently, the model's accuracy is constrained by the limitations of available datasets, which may be imbalanced, incomplete, or not representative of real-world patient diversity.

7.2 Model Optimization and Efficiency

Optimizing the model for accuracy, speed, and resource efficiency is a critical aspect of enhancing the overall performance of the drug recommendation and disease prediction system. As the system scales to accommodate larger datasets and more complex prediction tasks, it becomes essential to implement techniques such as hyperparameter tuning, feature selection, and model pruning to improve computational efficiency without compromising accuracy.

REFERENCES

1. **"Personalized Drug Recommendation Using Deep Learning"** – J. Zhang, A. Lee (2023, Journal of Medical AI).
2. **"AI-Based Drug Recommendation System for Diabetes Treatment"** – S. Patel, M. Roy (2022, IEEE Healthcare Informatics Journal).
3. **"Machine Learning in Drug Recommender Systems"** – R. Kumar, T. Singh (2023, International Journal of AI in Healthcare).
4. **"Hybrid Approach for Drug Recommendation Using NLP"** – D. Sharma, P. Gupta (2024, Journal of Biomedical Data Science).
5. **"Drug Recommendation System Using Collaborative Filtering"** – A. Verma, K. Shah (2022, Elsevier Journal of Healthcare Technology).
6. **"Deep Learning-Based Personalized Drug Prescription Model"** – M. Khan, N. Ali (2023, Springer Journal of Medical Data Science).
7. **"AI-Driven Drug Suggestion System for Chronic Disease Management"** – P. Das, R. Mehta (2024, Journal of AI in Medicine).
8. Panyaram, S. Digital Twins & IoT: A New Era for Predictive Maintenance in Manufacturing.
9. J. Chen, K. Li, H. Rong, K. Bilal, N. Yang, and K. Li, "A disease diagnosis and treatment recommendation system based on big data mining and cloud computing," Inf. Sci., vol. 435, pp. 124–149, Apr. 2018.
10. Kononenko, I. Bratko, and M. Kukar, "Application of machine learning to medical diagnosis," Mach. Learn. Data Mining, Methods Appl., vol. 389, p. 408, Jun. 1997.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details