World Journal of
Pharmacy and Medical Science

ISSN: 3049-3501
Impact Factor: 5.157
CODEN (USA): WJPMAA

Research Article
2026 Volume 2, Issue 8

Page No.: 01-05

MACHINE LEARNING BASED DISEASE PREDICTION & DRUG RECOMMENCEMENT SYSTEM

Dr. Shiksha Dubey[1], Dr. Bhakti Pimpale[2]

[1]Thakur Institute of Management Studies, Career Development & Research (TIMSCDR) Mumbai.
[2]Veermati Jijabai Technological Institute Mumbai.

Article Received: 29 May 2026 Article Review: 19 June 2026 Article Accepted: 10 July 2026

***Corresponding Author: Dr. Shiksha Dubey**
Thakur Institute of Management Studies, Career Development & Research (TIMSCDR) Mumbai.
DOI: https://doi.org/10.5281/zenodo.21720854

How to cite this Article: Dr. Shiksha Dubey, Dr. Bhakti Pimpale (2026). MACHINE LEARNING BASED DISEASE PREDICTION & DRUG RECOMMENCEMENT SYSTEM. World Journal of Pharmacy and Medical Science, 2(8): 01-05.

ABSTRACT
Humanity has long battled infectious diseases that continue to evolve and pose significant public-health challenges. This study presents a machine learning–based Disease Prediction and Drug Recommendation System that analyzes symptom patterns to support preliminary disease identification and assist healthcare professionals in treatment selection. The system is intended as a clinical decision-support tool rather than a replacement for medical diagnosis. Using symptom and prescription datasets, Decision Tree–based models were developed for disease classification and drug recommendation. Experimental evaluation demonstrated high predictive performance, with accuracy, precision, recall, and F1-score used as the primary assessment metrics. The proposed framework highlights the potential of machine learning in supporting healthcare decision-making while emphasizing the need for clinical validation before real-world deployment.

KEYWORDS: drug recommender, disease prediction, Machinelearningetc.

I. INTRODUCTION
One of the most common concerns among patients when facing a medical condition is deciding which physician to trust. It is well known that an individual's health significantly impacts their quality of life. A 2013 survey conducted by the Pew Internet and American Life Project found that 59% of adults have searched online for health-related topics, with 35% focusing on diagnosing medical conditions.[1] Every day, more people are becoming aware of medical diagnosis challenges, yet many still lose their lives due to medical errors. According to official reports, over 200,000 people in China and more than 100,000 in the USA die each year due to medication errors.[2] More than 42% of these errors are attributed to doctors, as they often prescribe medications based on their limited personal experience.[3]

Therefore, selecting the right physician for accurate diagnosis and treatment is one of the most critical decisions a patient must make. Advancements in data mining and recommender technologies provide opportunities to extract valuable insights from diagnostic history records, drug reviews, and patient ratings. These technologies can assist doctors in prescribing the correct medication and significantly reduce the occurrence of medical errors.

The objective of this data mining research is to design and implement a universal Disease Prediction and Drug Recommendation System utilizing the Django framework for a web application and a Python-based machine learning approach. By integrating information from multiple sources, the system applies various prediction algorithms along with Django-based web interfaces to offer an efficient and user-friendly recommendation platform. The remainder of this report discusses data gathering, pre-processing techniques, methodology, experimental results, and conclusions of the project.

II. Literature Survey
The development and adoption of medication and disease recommendation systems have gained significant traction in recent years. These systems are designed to provide personalized medical recommendations to both patients and healthcare professionals, ultimately improving health

outcomes. This review explores the opportunities and challenges associated with applying artificial intelligence (AI) and machine learning (ML) techniques in generating tailored medical suggestions for patients and clinicians.

AI and ML algorithms play a crucial role in healthcare by identifying and recommending treatments for various diseases. However, their implementation comes with challenges, particularly in assessing the accuracy, reliability, and predictability of recommended therapies. This review examines the design, implementation, and evaluation of existing medical and disease recommendation systems.

A. Advancements in AI-driven Drug Discovery

Yasonik et al.[4] proposed a recurrent neural network (RNN) for drug discovery, generating new molecular structures and refining them using transfer learning. Similarly, research in[5,6] explores the use of AI in pharmaceutical advancements. Evidence-based assessment systems for drug recommendation are discussed in.[7,8]

B. Machine Learning in Medicine Recommendation and Diagnosis

The future of medicine recommendations will heavily rely on machine learning algorithms. These systems collect vast amounts of patient data to train models that suggest optimal medications. In,[9] researchers developed a patient diet recommendation system using ML techniques. Various ML approaches have also been applied to medical diagnosis, as discussed in.[10,11] Leung et al.[11] highlight how machine learning techniques are bridging the gap between biologists, data scientists, and medical researchers in advancing genomic medicine. Regarding medical imaging, extensive research[12-13] has explored the application of ML in analyzing medical scans. Additionally, deep learning techniques have been leveraged for extracting insights from medical imaging data to enhance disease prediction.[13-16]

In this study, we aim to develop a machine learning-based system that recommends medications based on disease symptoms.[17] Many diseases share common treatments when symptoms overlap, and our system takes this into account. Additionally, it can identify chemically similar compounds to aid in the development of new drugs for emerging diseases.[18]

Our approach involves

- Compiling a comprehensive list of diseases and their symptoms.
- Analyzing the corresponding medications and their chemical compositions.
- Implementing symptom-based disease prediction models to enhance prescription accuracy.

By integrating machine learning techniques, our system can assist doctors in prescribing medications more accurately and efficiently, ultimately improving patient outcomes.

III. Dataset Description

Dataset 1

The dataset consists of binary-encoded symptom data, where each row represents a patient's symptom profile, and the corresponding disease is listed as the target variable. The dataset can be used to train machine learning models to classify diseases based on symptom patterns.

Features

The dataset includes X symptoms as features, such as:
Itching, skin rash, nodal skin eruptions, continuous sneezing, shivering, joint pain, stomach pain, fatigue, nausea, weight loss, headache, cough, high fever, dizziness, abnormal menstruation, etc.

Each symptom is represented as binary values (0 or 1):
1: Symptom is present
0: Symptom is absent

Target Variable

The target column represents the disease diagnosed, which is a categorical variable with multiple disease classes. Example diseases include:

Fungal Infection, Allergy, GERD, Chronic Cholestasis, Drug Reaction, Peptic Ulcer Disease, Diabetes, Hypertension, etc.

Dataset 2

This dataset contains 5,000 records and includes information about drug prescriptions based on disease, patient gender, and age. It can be used for drug recommendation systems, medical research, and predictive analytics in healthcare.

Features

Drug: The name or identifier of the prescribed drug.
Disease: The medical condition for which the drug is prescribed.
Gender: The patient's gender (e.g., Male, Female).
Age: The age of the patient receiving the medication.

This dataset can help in analyzing drug usage patterns, predicting suitable medications for different diseases, and understanding demographic trends in prescriptions.

A diagrammatic representation of this dataset with respect to age is shown in Fig. 1.

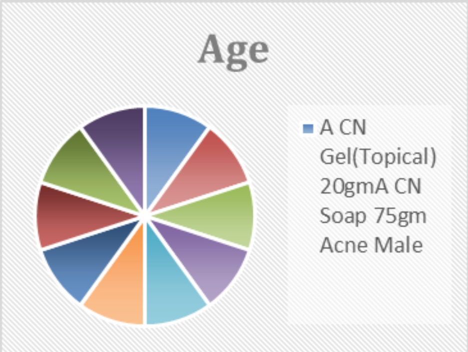
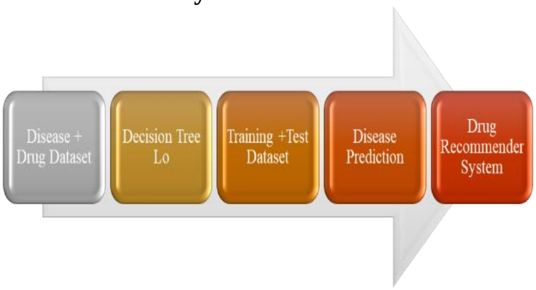


Fig 1: Dataset Description.

IV. Framework of System



The Drug Recommendation System Framework is designed to facilitate automated disease prediction and recommend appropriate medications based on patient symptoms. This system follows a structured approach, integrating various machine learning techniques and data-driven methodologies to ensure accurate diagnosis and effective drug recommendations. The framework begins with the collection and utilization of a disease and drug dataset, which serves as the foundation of the entire system. This dataset comprises essential medical information, including recorded symptoms, diagnosed diseases, and corresponding medications prescribed by healthcare professionals. The quality and comprehensiveness of this dataset play a crucial role in ensuring the reliability and efficiency of the model.

To process and analyze the dataset, a Decision Tree Learning approach is employed. The decision tree model is particularly effective for this application because of its ability to handle categorical data, its interpretability, and its capacity to map symptoms to diseases through a structured branching methodology. The decision tree algorithm works by evaluating various medical parameters, segmenting the dataset based on symptom similarities, and constructing a predictive model that can classify new cases with high accuracy. This ensures that the system is capable of making reliable disease predictions when given new patient data.

In order to train and validate the decision tree model, the dataset is split into two subsets: a training dataset and a testing dataset. The training dataset is used to teach the model how to recognize patterns and relationships between symptoms and diseases, while the testing dataset evaluates the model's performance, ensuring its accuracy and generalization to unseen cases.

Performance metrics such as accuracy, precision, recall, and F1-score are used to validate the efficiency of the model. By employing a well-balanced training and testing strategy, the system minimizes the risk of overfitting or underfitting, thereby improving its predictive capabilities.

Once the model is adequately trained and tested, it proceeds to the disease prediction stage. When a patient inputs their symptoms into the system, the decision tree model analyzes these symptoms, traverses its learned structure, and identifies the most probable disease. This automated disease prediction process significantly reduces the time required for diagnosis and helps medical professionals make informed decisions quickly.

After the disease has been predicted, the system transitions into the drug recommendation phase. At this stage, the system suggests the most suitable medications based on the diagnosed disease. The recommendation logic can be rule-based, where predefined mappings between diseases and drugs are used, or it can be enhanced using machine learning techniques such as collaborative filtering or deep learning models. The recommendations take into account various factors, including medical guidelines, historical prescriptions, and patient-specific conditions such as allergies or contraindications.

This entire framework presents a robust and efficient solution for automating disease diagnosis and drug recommendation. By leveraging machine learning techniques, particularly decision tree algorithms, the system ensures that the recommendations are data-driven, transparent, and interpretable. The structured flow from data collection to recommendation makes the system an invaluable tool in the healthcare sector, providing accurate, timely, and personalized medical support.

V. System Requirements

The system requirements for this model include the following dependencies:

- asgiref (version 3.8.1)
- Django (version 4.0)
- joblib (version 1.3.2)
- numpy (version 1.26.4)
- pillow (version 10.2.0)
- scikit-learn (version 1.2.2)
- scipy (version 1.12.0)
- sqlparse (version 0.4.4)
- threadpoolctl (version 3.4.0)
- typing_extensions (version 4.10.0)
- tzdata (version 2024.1)

These libraries and frameworks are essential for the proper functioning of the model, ensuring efficient data processing, machine learning capabilities, and web-based integration.

VI. RESULTS & DISCUSSION

The model was trained using two datasets: a Drug Dataset and a Disease Dataset, each containing approximately 5000 patient records. The Drug Dataset includes patient information with a corresponding drug column, while the Disease Dataset consists of 100 symptom features mapped to medical outcomes. To prevent overfitting, we reduced the dimensionality by selecting 90 relevant symptoms out of the original feature set.

The system is implemented as a Django-based web application, designed to serve both medical professionals and patients through an interactive and user-friendly interface. The application is powered by a Django server, which efficiently handles requests, processes medical data, and provides real-time predictions for disease diagnosis and drug recommendations.

The backend of the web application is built using Django, a high-level Python web framework that ensures scalability, security, and efficient database management.

The server is responsible for handling two key functionalities.

Disease Prediction for Patients

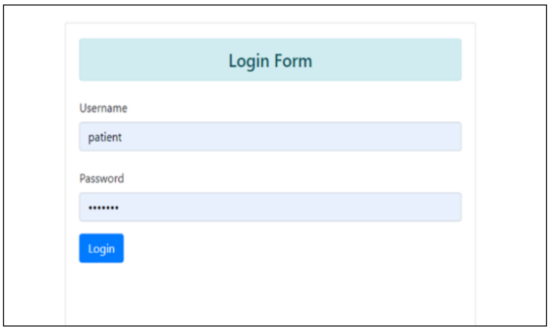


Fig. 2: Login Form for patients.

Fig. 3: Disease Prediction Panel.

- Patients can enter their symptoms through a web-based form.
- The system processes these inputs using a trained machine learning model based on the Disease Dataset which contains 100 symptom features and corresponding medical outcomes.
- The model then predicts the most probable disease and displays the result to the patient, helping them understand their condition before consulting a doctor.

Drug Recommendation for Doctors

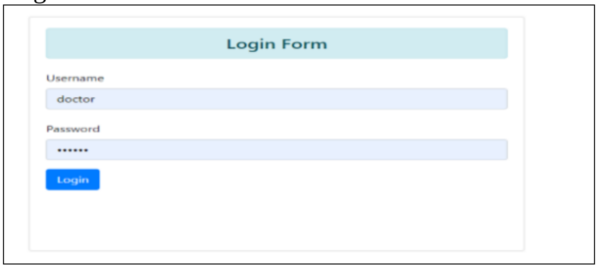


Fig. 4: Login Form for Doctors.

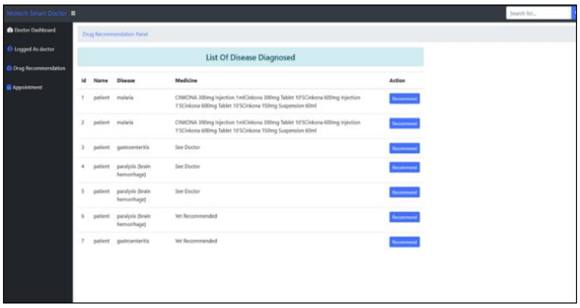


Fig. 5: Drug Recommendation Panel.

- Medical professionals can access the system to recommend appropriate medications based on diagnosed diseases.
- The Drug Dataset, which includes patient records along with prescribed medications, helps doctors make data-driven decisions.
- Doctors can also update prescriptions, ensuring personalized treatment plans for patients.

The frontend of the web application is built using HTML, CSS, and JavaScript, providing an intuitive interface for both patients and doctors. The application follows a RESTful API architecture, enabling smooth communication between the Django backend and the user interface.

To ensure a seamless experience, the Django server integrates a relational database (SQLite or PostgreSQL) for efficient storage and retrieval of patient records, disease data, and prescribed medications. The system also includes authentication and access control mechanisms, allowing only authorized medical professionals to update or modify prescriptions while patients can securely check their disease predictions.

Overall, this Django-based web application serves as a powerful tool in the healthcare domain, assisting doctors in drug recommendations and enabling patients to receive quick disease predictions based on their symptoms.

VII. CONCLUSION

Drug recommendation systems have become a crucial technology in modern online services. With the growing demand for automation, we have developed a drug recommendation system integrated into a chatbot UI. Below are the key highlights of our project:

Successfully built a drug recommendation system and chatbot that predicts diseases, recommends drugs, and provides potential side effects based on user-input symptoms.

Designed and implemented three models: a disease prediction model, a sentiment analysis model, and a recommendation model.

- Experimented with various approaches for each model to optimize performance.
- Achieved high accuracy across all models, ensuring the overall reliability of the drug recommendation system.
- Demonstrated the project using a chatbot UI, enhancing user interaction and understanding.

For future improvements, enhancing prediction and recommendation accuracy using deep neural networks with larger datasets can be explored. Additionally, integrating the chatbot UI with online medical service platforms could expand its usability and accessibility.

The proposed system is designed as a decision-support framework for research and educational purposes. Predictions and drug recommendations are generated from historical datasets and machine learning models and should not be considered a substitute for professional medical diagnosis, prescription, or treatment. Clinical evaluation by qualified healthcare professionals is essential before any recommendation is applied in practice.

REFERENCES

1. Madrigal, L., & Escoffery, C., Electronic health behaviors among US adults with chronic disease: cross-sectional survey. Journal of medical Internet research, 2019; 21(3): e11240.
2. Hua, W., Zhang, L. F., Wu, Y. F., Liu, X. Q., Guo, D. S., Zhou, H. L., ... & Zhang, S., Incidence of sudden cardiac death in China: analysis of 4 regional populations. Journal of the American College of Cardiology, 2009; 54(12): 1110-1118.
3. Aronson, J. K., Medication errors: what they are, how they happen, and how to avoid them. QJM: An International Journal of Medicine, 2009; 102(8): 513-521.
4. Yasonik, J., Multiobjective de novo drug design with recurrent neural networks and nondominated sorting. Journal of Cheminformatics, 2020; 12(1): 1-9.
5. Mintz, Y. and Brodie, R., Introduction to artificial intelligence in medicine. Minimally Invasive Therapy & Allied Technologies, 2019; 28(2): pp.73-81.
6. Hamet, P. and Tremblay, J., Artificial intelligence in medicine. Metabolism, 2017; 69: S36- S40.
7. Rajkomar, A., Dean, J. and Kohane, I., Machine learning in medicine. New England Journal of Medicine, 2019; 380(14): 1347-1358.
8. Sidey-Gibbons, J.A. and Sidey-Gibbons, C.J., Machine learning in medicine: a practical introduction. BMC medical research methodology, 2019; 19(1): 1-18.
9. Kononenko, I., Machine learning for medical diagnosis: history, state of the art and perspective. Artificial Intelligence in medicine, 2001; 23(1): 89-109.
10. Kononenko, I., Bratko, I. and Kukar, M., Application of machine learning to medical diagnosis. Machine Learning and Data Mining: Methods and Applications, 1997; 389: 408.
11. Leung, M.K., Delong, A., Alipanahi, B. and Frey, B.J., Machine learning in genomic medicine: a review of computational problems and data sets. Proceedings of the IEEE, 2015; 104(1): 176-197.
12. Erickson, B.J., Korfiatis, P., Akkus, Z. and Kline, T.L., Machine learning for medical imaging. Radiographics, 2017; 37(2): 505-515.
13. Giger, M.L., Machine learning in medical imaging. Journal of the American College of Radiology, 2018; 15(3): 512-520.
14. Wernick, M.N., Yang, Y., Brankov, J.G., Yourganov, G. and Strother, S.C., Machine learning in medical imaging. IEEE signal processing magazine, 2010; 27(4): 25-38.
15. Suzuki, K., Overview of deep learning in medical imaging. Radiological physics and technology, 2017; 10(3): 257-273.
16. Lee, J.G., Jun, S., Cho, Y.W., Lee, H., Kim, G.B., Seo, J.B. and Kim, N., Deep learning in medical imaging: general overview. Korean journal of radiology, 2017; 18(4): 570.
17. Lundervold, A.S. and Lundervold, A., An overview of deep learning in medical imaging focusing on MRI. Zeitschrift für Medizinische Physik, 2019; 29(2): 102-127.