Diagnosis Of Disease Using Machine Learning

Shankar Adkile, Kashinath Andewadi, Atik Misal Somnath Lambe, Priyanka Kulkarni

E&TC Student of KIT Shelve, Assistant Professor at KIT

[1]*Received: 18/04/2025; Accepted: 16/05/2025; Published: 22/05/2025*

Abstract

Health is a fundamental aspect of a person's life, and understanding the details of various diseases is crucial for everyone. Today, many diseases are commonly encountered, and both individuals and healthcare professionals need to accurately diagnose these conditions. This paper presents a fever classification and remedial recommendation system using the decision tree algorithm, while also outlining the basic characteristics of decision trees. In medical decision-making (such as classification and diagnosis), there are numerous situations where decisions must be made effectively and reliably. Simple decision-making models with the ability to learn automatically are ideal for handling such tasks. Decision trees are a trustworthy and efficient technique for decision-making, offering high classification accuracy and a straightforward representation of acquired knowledge. They have been successfully applied in various areas of medical decision-making.

1. Introduction

The rapid spread of infectious diseases, especially those characterized by fever and related symptoms like chills, headaches, and runny nose, has highlighted the urgent need for efficient public health monitoring. Fever detection has become a crucial tool for early disease identification and prevention, particularly in high-traffic areas such as hospitals, airports, and schools. This project aims to develop a system for monitoring fever detection using Python and machine learning, specifically leveraging the Decision Tree Algorithm. The system is designed to accurately assess symptoms in individuals, contributing to early disease detection and intervention. The significance of this project lies in its potential to enhance public health safety by offering a reliable, scalable solution for fever detection. By using machine learning algorithms, the system improves detection accuracy and provides a scalable approach that can be applied in diverse settings. Early identification of fever can play a pivotal role in preventing the spread of infectious diseases and safeguarding public health. Disease diagnosis is a core aspect of healthcare, typically involving the identification of medical conditions through patient symptoms, history, and clinical findings. Traditionally, this process has relied on healthcare professionals' expertise, patient interviews, and diagnostic tests. However, symptoms are often vague and overlap across multiple diseases, which can make accurate diagnosis difficult, especially for complex or rare conditions. In recent years, Machine Learning (ML), a branch of artificial intelligence (AI), has become an essential tool in medical diagnosis by analyzing patterns in large datasets.ML algorithms can learn from historical medical data, symptoms, and outcomes, allowing them to predict specific diseases with high accuracy. This data-driven approach enables healthcare providers to make more informed decisions, detect diseases earlier, and recommend personalized treatment options. Machine learning models can be trained on extensive datasets containing symptoms, demographics, lab results, and medical histories, allowing them to recognize complex patterns that may be challenging for humans to identify. Objectives Develop a machine learning-based predictive model capable of accurately identifying diseases based on a set of input symptoms. This model aims to facilitate early diagnosis and enable faster intervention, potentially reducing the need for physical temperature measurements in certain cases. Create machine learning tools to assist healthcare providers by suggesting possible conditions based on clinical data. Develop algorithms that detect diseases

[1] *How to cite the article:* Adkile S., Andewadi K., Lambe A.M.S., Kulkarni P. (May, 2025); Diagnosis Of Disease Using Machine Learning; *International Journal of Advances in Engineering Research*, Vol 29, Issue 5, 43-47

in their early stages by analyzing patients' symptoms. Build machine learning models that provide personalized risk assessments and diagnoses based on an individual's unique symptom profile.

2. Objectives

Develop a machine learning based predictive model capable of accurately identifying diseases based on a set of input symptoms. This model aims to facilitate early diagnosis and enable faster intervention, potentially reducing the need for physical temperature measurements in certain cases. Create machine learning tools to assist healthcare providers by suggesting possible conditions based on clinical data. Develop algorithms that detect diseases in their early stages by analyzing patients' symptoms. Build machine learning models that provide personalized risk assessments and diagnoses based on an individual's unique symptom profile.

3. Proposed Work

Data Collection:

The first step in this project is collecting a comprehensive dataset that includes:

- Symptoms: A list of symptoms reported by patients (e.g., fever, cough, headache, etc.)
- Diseases: The diseases associated with those symptoms (e.g., flu, COVID-19, malaria, etc.)
- Patient Demographics: Additional information such as age, gender, and medical history, which may impact diagnosis.

Sources of data may include:

- Open medical datasets (e.g., Disease-Symptom Database, UCI Machine Learning Repository)
- Clinical records (with proper privacy and ethics considerations)

Data Preprocessing:

- Cleaning: Handle missing values, remove duplicates, and standardize symptom names.
- Encoding: Convert categorical data (e.g., symptoms, disease labels) into numerical format using techniques like one-hot encoding or label encoding.
- Normalization: Normalize numerical data (e.g., age, severity levels) for consistency.

The proposed project seeks to develop a machine learning-based system for disease diagnosis using patient symptoms. By automating the diagnostic process, the system aims to support healthcare professionals in making quicker and more accurate diagnoses, ultimately enhancing patient outcomes and improving overall healthcare efficiency.

4. System Design

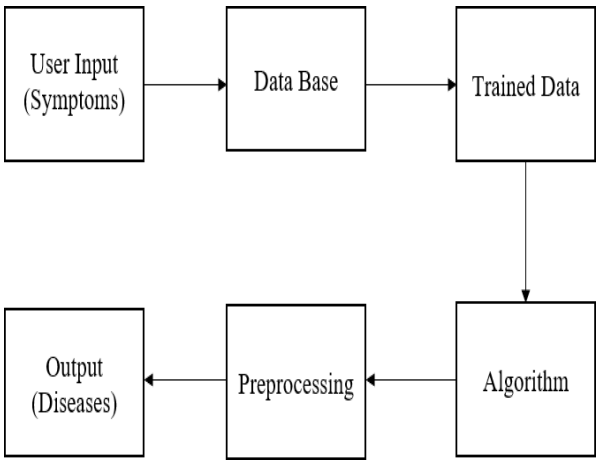


Fig. 1 Block Diagram of System Design

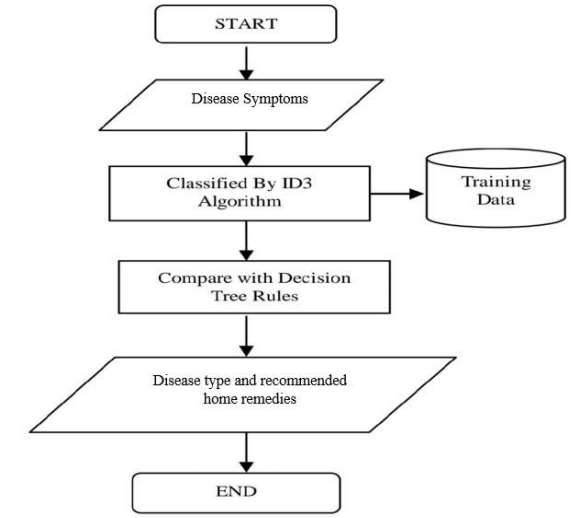


Fig. 2 Flow Chart of Decision Tree Algorithm

User Input (Symptoms): The inputs are given as a Symptoms like (e.g. pus filled pimples, blackheads, yellow urine, runnynose).

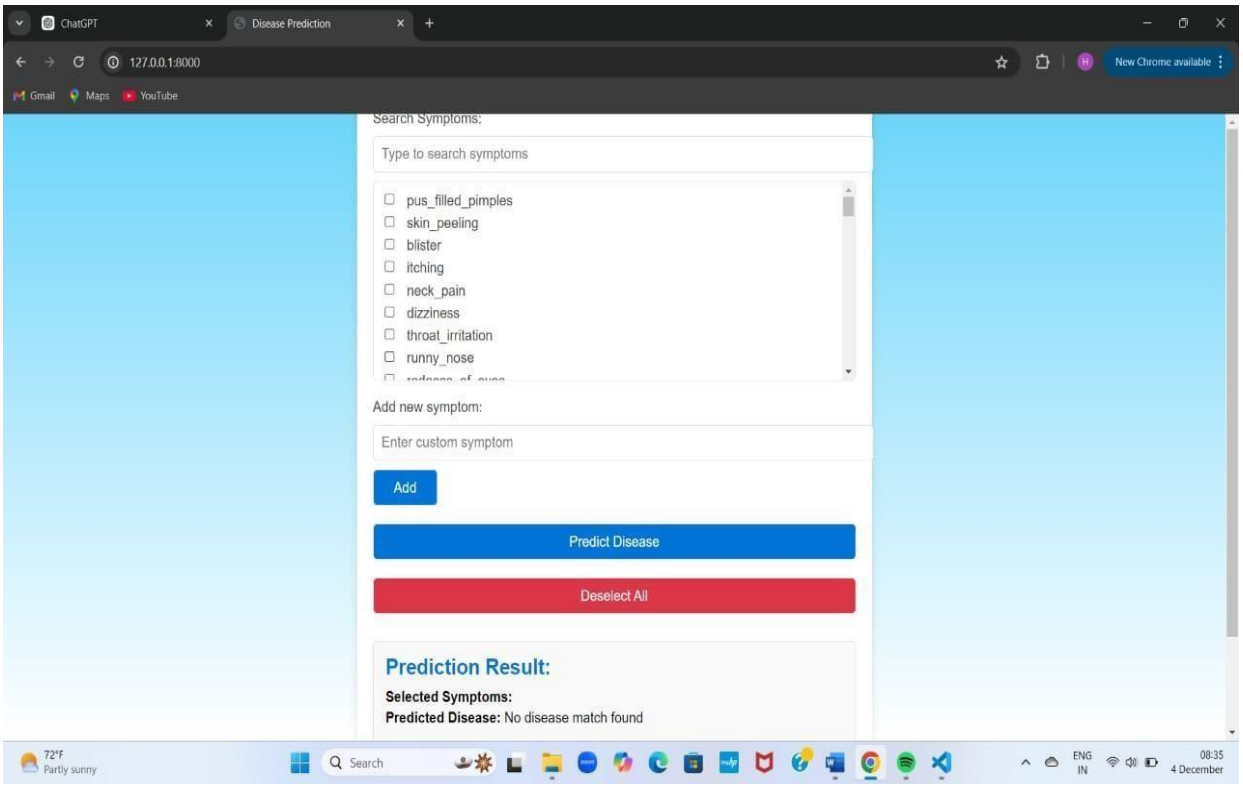
Data Base: This block represents the interaction point for end user or external applications to interact with the database.

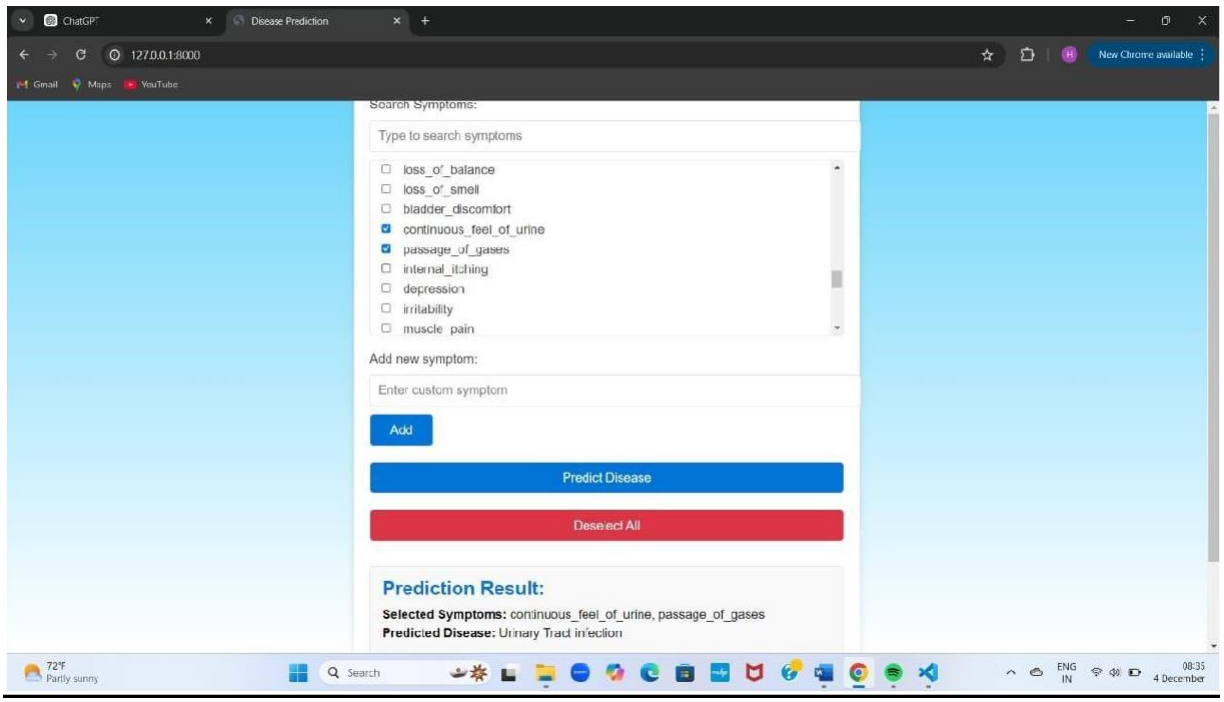
Algorithm: Algorithm is a step by step process where each term is classified by the given inputs.

Preprocessing: The user input is sent to a processing block, where it is interpreted or manipulated.

Training Data: Create and train the decision tree classifier using the training data. **Output (Diseases):** After processing, the system generates an output, which could be in terms of disease.

5. Results


ChatGPT
Disease Prediction
127.0.0.1:8000
Search Symptoms:
Type to search symptoms
pus_filled_pimples
skin_peeling
blister
itching
neck_pain
dizziness
throat_irritation
runny_nose
Add new symptom:
Enter custom symptom
Add
Predict Disease
Deselect All
Prediction Result:
Selected Symptoms:
Predicted Disease: No disease match found


ChatGPT
Disease Prediction
127.0.0.1:8000
Search Symptoms:
Type to search symptoms
loss_of_balance
loss_of_smell
bladder_discomfort
continuous_feel_of_urine
passage_of_gases
internal_itching
depression
irritability
muscle_pain
Add new symptom:
Enter custom symptom
Add
Predict Disease
Deselect All
Prediction Result:
Selected Symptoms: continuous_feel_of_urine, passage_of_gases
Predicted Disease: Urinary Tract Infection

6. Conclusion

In conclusion, the use of machine learning (ML) techniques for diagnosing diseases based on symptoms has proven to be highly promising, providing an efficient, scalable, and often more accurate alternative to traditional diagnostic methods. By analyzing large datasets that map symptom-disease relationships, ML models can uncover patterns and correlations that may be overlooked by human experts.ML has the potential to transform disease diagnosis by leveraging patient symptoms and medical data. With the use of algorithms like decision trees, neural networks, and support vector machines, healthcare systems can achieve quicker, more precise diagnoses, particularly in complex cases where symptoms are ambiguous or overlap. These models can process vast amounts of data, identify key patterns, and predict the likelihood of specific diseases based on a combination of symptoms, medical history, and other diagnostic factors. However, challenges still exist, such as the need for high- quality labeled data, addressing biases in training datasets, and ensuring the interpretability of complex models. Despite these hurdles, the integration of ML in disease diagnosis offers great promise, especially in resource-constrained settings where access to healthcare professionals and diagnostic tests is limited. As the field continues to advance, future improvements in ML-based diagnostic tools, along with greater collaboration between healthcare providers, data scientists, and researchers, will likely lead to more efficient, cost-effective, and accessible healthcare solutions for diagnosing a wide range of diseases, ultimately enhancing patient outcomes globally.

7. Conflict of Interest

The authors declare that they have no conflict of interest.

8. Funding Declaration

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

9. References

Alanazi, R. (2022). Identification and prediction of chronic diseases using machine learning approach. *Journal of Healthcare Engineering, 2022*, 2826127. https://doi.org/10.1155/2022/2826127

Gaurav, K., Kumar, A., Singh, P., Kumari, A., Kasar, M., & Suryawanshi, T. (2023, June). Human disease prediction using machine learning techniques and real-life parameters. *International Journal of Engineering, 36*(6), 1092–1098. https://doi.org/10.5829/ije.2023.36.06b.01

Gomathy, C. K., & Naidu, A. R. (2021, October). The prediction of disease using machine learning. *International Journal of Scientific Research in Engineering and Management, 5*(10). https://doi.org/10.55041/IJSREM17607

About Author

Mr. Shankar Shivraj Adkile is currently pursuing a Bachelor's Degree in Electronics and Telecommunication Engineering at Dr. Babasaheb Ambedkar Technological University. His academic and research interests include machine learning, artificial intelligence, and healthcare technologies.