AI-Powered Symptoms-Based Disease Detection and Medicinal Plant Identification System

Raveesh R[1], Shruthi V S[2], Rohan Rao G H[3], Vaishnavi K[4], Varnna Balakrishnan[5]
[1,3,4,5]*Student, Srinivas Institute of Technology (S.I.T), Mangalore, Karnataka, India*
[2]*Assistant Professor, Srinivas Institute of Technology (S.I.T), Mangalore, Karnataka, India*
doi.org/10.64643/IJIRTV12I12-206768-459

***Abstract*— In today's fast-paced world, people often rush to modern healthcare for even minor health issues, leading to rising costs and the slow fading of traditional healing knowledge passed down through generations. This project addresses that growing concern by offering a simple yet intelligent approach to early disease detection and natural treatment guidance.**

What we built is a web platform that approaches the problem from two angles. First, it takes the symptoms a user describes, and runs them through machine learning classifiers to suggest possible conditions, along with a few precautions the person can take right away. Second, it also lets people photograph a plant and then get an identification back, but paired with details on how that same plant has been used in medicinal ways. And a chatbot sits on top of both parts, asking follow up questions in plain language so the whole output never feels like you are reading a diagnostic report, not even a little. The goal was never to replace a doctor, it was more like to give people a kind of starting point, you know, something that helps them figure out if a situation really needs professional attention or if it can reasonably be handled at home, at the same time as keeping traditional healing knowledge reachable for anyone who is curious enough to look.

***Index Terms*—Disease prediction, medicinal plants, machine learning, deep learning, CNN, health care systems, symptom analysis, chatbot.**

I. INTRODUCTION

People today depend more on clinics and medicines for everyday health issues, while the traditional knowledge of medicinal plants and at home remedies is slowly fading out. As this dependence keeps growing it has also driven healthcare costs up, and it reduced the awareness people had about natural healing practices that used to pass from one generation to the next.

This project tries to help with both sides, kinda at once, by guessing possible diseases from symptoms with machine learning methods such as Decision Tree, Random Forest, and KNN. It also works on recognizing medicinal plants from images someone uploads, using deep learning models like CNN and MobileNetV2, trained using medicinal plant datasets.

The platform offers an intuitive web-based interface allowing users to input their symptoms or submit plant photographs for immediate feedback. An integrated chatbot explains di- agnostic results, recommends natural treatments, and offers safety guidelines using straightforward language. Furthermore, educational materials including YouTube resources help users learn proper medicinal plant application methods.

This project sort of mixes current machine learning with more traditional plant-based medicine, so it becomes practical, and easy to use, as a kind of healthcare support system. It lets users quickly check their health status, and also look into natural remedies without wasting time scrolling through unreliable sources.

II. RELATED WORK

Disease prediction together with medicinal plant identification has really been researched for quite a while now. In the beginning, older healthcare systems mostly relied on rule-based approaches that were made by medical specialists, but honestly these things were hard to revise, and they tended to be less capable when the situation was unusual or when datasets got really big.

Machine learning changed the approach pretty a lot. Researchers started training KNN, Decision Tree and Random Forest classifiers straight on symptom records, instead of using manually coded rules, you know. KNN is kind of intuitive: it looks for the nearest

similar cases in the training data, but the same idea is what makes it sluggish on big datasets because each prediction kind of has to compare against the full training set. Random Forest kinda helps around that by averaging across a bunch of trees, and as a result it becomes more resilient to noisy or tangled symptom combinations. Both of these approaches have an advantage too, they give outputs that are fairly easy to interpret, which is important in medical settings.

As deep learning kept growing, more powerful approaches started popping up, especially for image driven work such as medicinal plant identification. Convolutional Neural Networks showed up everywhere, mostly because they can pull out essential visual features and then sort the plants with a pretty high level of accuracy. After that, more developed designs like MobileNetV2 and VGG16 took things further, they also managed to produce impressive outcomes while still holding computational demands at a relatively low level, which is kind of the goal in practice.

These frameworks prove invaluable in practical scenarios requiring image analysis capabilities.

Recent healthcare platforms are starting to lean on chatbot interfaces as well, to make predictions feel more or less understandable. Rather than just throwing out a diagnosis on screen, these systems guide the users by walking them through the outcome in plain words. They also handle follow up questions, so the whole healthcare support side becomes more reachable for people who dont really have medical knowledge.

In this project, we bring together machine learning based disease prediction, deep learning driven medicinal plant identification, plus chatbot assistance, all inside one single platform that is meant to be easy on the eyes and easy to use. The intention is to deliver dependable results in a way thats simple, and still easy to grasp, so basically anyone can use the system without needing technical know-how or medical expertise.

III. PROBLEM STATEMENT

Today, a lot of people rely on hospitals and medicines even for everyday health issues like fever or cold. This situation has made healthcare costs much higher, medical facilities get overcrowded, and medicines get used too often, while the traditional know how about checking early symptoms and using small home remedies slowly fades away.

The tools currently available don't help much either, or at least not in a way that feels useful. Most health platforms are built around diagnosis and they stop right there, giving almost no practical advice on what a person is supposed to do with that information, and then you're kinda stuck. Navigation is another thing many of these systems were clearly made with a technically literate user in mind, so if you're outside that group you quickly get lost, like, completely. There's also this noticeable gap in platforms that treat modern disease prediction and traditional plant-based medicine as two sides of the same coin, instead of totally separate concerns. Folks who could benefit from both get pushed into looking in different places, if they look at all.

What we set out to build was a single place that actually addresses those gaps. A user should be able to describe what they are feeling, get a response that is clear and doable, and if they are curious about a plant-based remedy, photograph a plant and have it identified too, all of that without needing any medical training or technical background to understand the results. The chatbot piece exists for that last part; it takes what the model outputs and turns it into plain language so nothing gets muddled between the prediction and the person reading it.

IV. PROPOSED SYSTEM

The platform we built runs on a combination of machine learning and deep learning models working together under a single web interface. This Artificial Intelligence based healthcare platform uses machine learning and deep learning to figure out what disease someone has and to identify plants that can be used as medicine. Users can interact with the platform in two ways, either describing their symptoms to get a probable diagnosis, or uploading a photo of a plant to have it identified. Having both features in one place means someone can check a symptom, get a suggestion that includes an herbal remedy, and then immediately confirm the plant, without having to switch between tools.

Before any data gets to a model, it passes through some kind of preparation phase. At that stage, symptom inputs are cleaned up, transformed into a representation that the classifiers can actually use, and

normalized too. For images, there is resizing plus pixel normalization, so the deep learning models end up seeing a consistent sort of input even when the photo came from different hardware or devices. For the disease prediction part, it runs via Decision Tree, Random Forest, and KNN, while plant identification is tackled using CNN and MobileNetV2.

Once everything is processed the platform gives results that are somehow useful, like the predicted disease or plant name, the confidence score, and some practical suggestions such as precautions and possible remedies. it also has educational info for people who want to learn more about the issue, or medicinal plant, in a deeper sense.

A chatbot kinda ties everything together at the end. Instead of just leaving users to figure out what a prediction means on their own, it translates the outcome into plain language, also it takes care of follow up questions without making the person re-enter anything. Basically, the whole experience runs through a straightforward web site, no installation, no account needed so there's hardly any gap between having a health concern and getting a useful first response, as small as we can make it.

A. Algorithm Description and Workflow

When someone opens the platform, they either type in their symptoms or upload a plant photo, that choice kinda decides what pipeline runs. the input gets cleaned and prepared first, then sent to the set of models that fits best, like ML classifiers when it's about symptoms, and deep learning models for images.

The thing that comes back includes the predicted illness or plant, a confidence score, and those concrete suggestions the user can actually do. Then the chatbot just continues from there, kinda guiding the person along through what the result really means in everyday terms, instead of leaving them alone to interpret just a label and a percentage by themselves.

V. METHODOLOGY

The way we do things is honestly pretty straightforward, but not in a boring way. We follow a process that involves collecting data, then getting it ready training our models. After that we make predictions, and that is basically it. Our system can handle two kinds of inputs: information about symptoms to predict diseases and pictures of plants to identify them.

The system uses verified public datasets for symptom ↔ disease records and labeled medicinal plant images. Symptom data got converted into binary feature vectors or so, while the plant images were resized and then normalized, for that reason to keep the model training consistent across the board.

For disease prediction we used a Decision Tree, Random Forest, and also KNN, as classifiers. Then for medicinal plant identification we switched to CNN and MobileNetV2 style models, basically those vision networks. All of them were evaluated with accuracy, precision, recall, and the F1-score so the results stay dependable across the different classes.

The platform gives results alongside confidence scores, precautions, and practical advice, rather than just showing raw predictions. You get a kind of simple user interface plus a chatbot support that helps users grok the outcomes better and then make more informed decisions, without getting lost.

A. Dataset and Feature Analysis

We use datasets that have information about symptoms and pictures of plants that are labeled. We look at the data to find the important features and we normalize and encode it to make our models work better.

B. Data Collection

We get our data from datasets that have information about symptoms and diseases and pictures of plants with labels.

C. Preprocessing

We clean and transform the data. We. Normalize the symp- tom data and resize and process the picture data.

D. Model Training

We use models like Decision Tree, Random Forest, KNN, CNN and MobileNetV2. We train them using supervised learning.

E. Prediction

Our system predicts what disease someone might have or what plant is in a picture, based on the input and gives us a confidence score along with some advice, on what to do and what to avoid.

The system uses disease prediction and plant identification models to give us the results.

VI. IMPLEMENTATION DETAILS

We built our system using a design. This makes it easy to add features and handle more users. Python is the program-ming language we used.

- We used libraries like Scikit-learn for machine learning tasks.
- TensorFlow and Keras are used for deep learning tasks.
- OpenCV helps us process images. We created two types of models.

Disease prediction models include Decision Tree, Random Forest and KNN. We trained these models on datasets of symptoms. We used learning for this. Plant identification mod-els use CNN and MobileNetV2. We trained these on labeled images of plants. We adjusted the model settings to get the results. When users interact with our system, they can input symptoms. Upload pictures of plants. The system prepares this input. Feeds it to the trained models. The models then make predictions. We show the results, including the disease or plant name and how confident we are on a website. The website is built with HTML, CSS and JavaScript. Flask handles the behind-the-scenes work. We store user data and prediction history in a database. We also have a chatbot. It explains the results to users. It suggests remedies. It guides users through the system. Our goal was to make the system easy to use and interactive. We want it to be helpful, in real-world healthcare situations.

VII. SYSTEM ARCHITECTURE

The system has four parts: Data Collection, Data Prepro- cessing, Model Processing and Presentation. This makes sure the system works well and gives results.

Data Collection Layer:

People using the system can enter symptoms to find out what disease they might have or upload pictures of plants to identify them.

Data Preprocessing Layer:

The information that people enter is cleaned up. Made ready to use. If someone enters symptoms this information is changed into a code. Made standard. If someone uploads a picture of a plant the picture is made bigger or smaller. Improved so the system can look at it better.

Model Processing Layer:

The system uses computer programs like Decision Tree, Random Forest and KNN to figure out what disease someone might have. It uses programs like CNN and MobileNetV2 to identify plants. These programs look at the information. Make predictions and they also say how sure they are about these predictions.

Presentation Layer:

The results are shown to people using the system through a website that's easy to use. This website shows what the system thinks might be wrong what people can do to feel and what they should be careful about. There is also a chatbot that helps people understand what the results mean and answers any questions they might have.

Overall, these parts of the system work together to give people fast, accurate and easy to use help with their health.

A. Architecture Workflow Explanation

Here is how the system works: first people using the system enter information like symptoms or a picture of a plant. Then this information is cleaned up. Made ready to use. Next special computer programs look at this information. Make predictions, about what might be wrong. Finally, the results are shown to people through the website. The chatbot helps them understand what everything means so they can get the help they need.

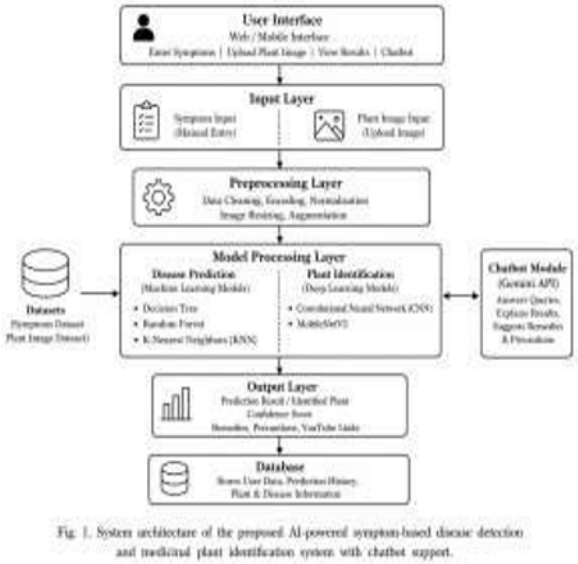

Fig. 1. System architecture of the proposed AI-powered symptom-based disease detection and medicinal plant identification system with chatbot support.

Fig. 1. System Architecture Diagram

VIII. RESULTS AND DISCUSSION

The system was tested using both a prepared dataset and actual user input to confirm it works accurately and reliably. The data was split into training and testing portions, and performance was measured using metrics like accuracy, precision, recall, and the F1-score.

Results showed that all disease prediction models performed well, with Random Forest outperforming Decision Tree and KNN mainly because it handles unseen data better and is less likely to overfit. For plant identification, CNN and MobileNetV2 both proved highly effective at recognizing plants from uploaded images.

The system makes few mistakes and when it does it is usually because of something called overfitting. The F1-score shows that the system is very good at finding a balance between saying something is wrong when it is not and saying something is right when it is not. When we looked at the mistakes the system made, we saw that KNN made a few mistakes than Random Forest and the deep learning models.

The system is very good at making predictions accurately. It is also very reliable. The system uses machine learning and deep learning. It has a user-friendly interface and a chatbot, which makes it easy to use and good for real healthcare situ- ations. The system is fast and accurate. People can depend on it. The disease prediction models and the plant identification models work well together to make the system a good tool, for healthcare.

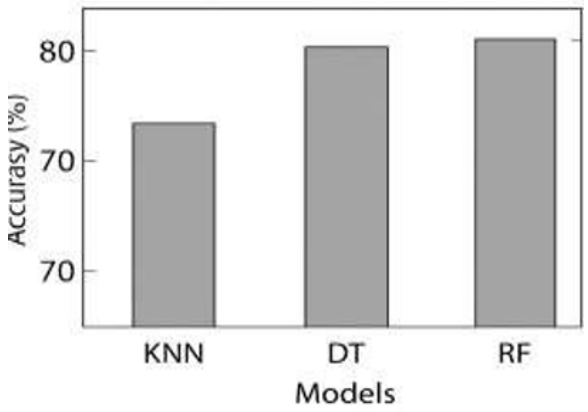

Fig. 2. Model Accuracy Comparison for Disease Detection

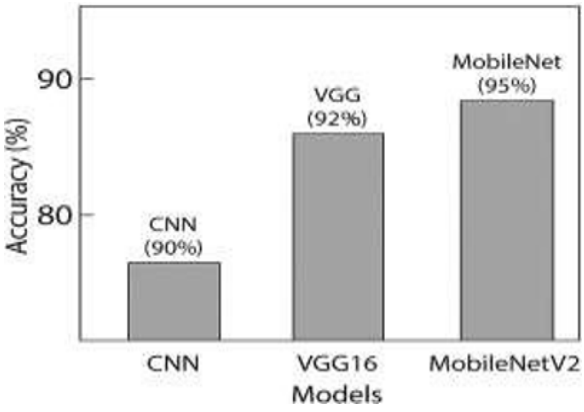

Fig. 3. Accuracy Comparison for Plant Identification Models

TABLE I
PERFORMANCE COMPARISON OF DISEASE PREDICTION MODELS

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Decision Tree	90.1	89.5	90.0	89.7
Random Forest	92.7	92.4	92.3	92.3
K-Nearest Neighbors (KNN)	88.3	87.6	88.0	87.8

Fig. 4. Performance Metrics Table

IX. CONCLUSION

This project demonstrates that an AI-powered system can effectively analyze user-input symptoms and plant images to provide accurate disease predictions and medicinal plant identification. By using machine learning and deep learning models, the system is able to deliver quick and reliable results, reducing the need for immediate medical consultation for minor health issues. The integration of a user-friendly interface and chatbot support further improves accessibility and helps users understand the results easily.

The system delivers solid accuracy and provides helpful outputs like precautions, remedies, and medicinal plant guidance, while also encouraging a valuable blend of modern and traditional healthcare. Going forward, it can be improved by adding more diseases, refining datasets, introducing multilingual support, and launching it as a mobile or cloud-based application to reach a much wider audience.

ACKNOWLEDGMENT

The authors want to thank Professor Shruthi V S, Assistant Professor of the Department of Computer

Science and Engi- neering, for providing them with the guidance, encouragement, and technical support they needed during the development of this work. The authors would like to extend their gratitude to the Department of Computer Science and Engineering for its support in providing the facilities required to complete this research.

REFERENCES

[1] R. Sood and V. Sharma, "Symptom Based Disease Prediction Using Machine Learning," 2024.
[2] K. Murali, "Medicinal Plant Identification," 2014.
[3] S. Gholap, "AI-Driven Symptom Analysis and Disease Prediction Assistant for Healthcare," 2025.
[4] S. Yadav, H. Sehrawat, Y. Singh, and V. Jaglan, "Machine Learning Approaches for Disease Prediction: A Review," 2022.
[5] Y. You and X. Gui, "Self-Diagnosis through AI-Enabled Chatbot-Based Symptom Checkers: User Experiences and Design Considerations," 2021.
[6] H. Wang, W.-C. Kwan, K.-F. Wong, and Y. Zheng, "CoAD: Automatic Diagnosis Through Symptom and Disease Collaborative Generation," 2023.
[7] W. He, X. Mao, C. Ma, Y. Huang, J. M. Hernández-Lobato, and T. Chen, "BSODA: A Bipartite Scalable Framework for Online Disease Diagnosis," 2022.
[8] L. Aissaoui Ferhi et al., "Enhancing Diagnostic Accuracy in Symptom-Based Health Checkers," 2024.
[9] K. Mulugeta, D. P. Sharma, and A. H. Mesfin, "Deep Learning for Medicinal Plant Species Classification and Recognition: A Systematic Review," 2023.
[10] L. Grinblat, L. C. Uzal, M. G. Larese, and P. M. Granitto, "Deep Learning for Plant Identification Using Vein Morphological Patterns," 2016.
[11] K. L. D. Viet et al., "Medicinal Plants Identification Using Federated Deep Learning," 2024.
[12] Fuster-Palà et al., "Optimized Machine Learning Classifiers for Symptom-Based Prediction," 2024.
[13] P. Dalvi, D. R. Kalbande, S. S. Rathod, H. Dalvi, and A. Agarwal, "Multi-Attribute Deep CNN for Medicinal Plant Detection," 2025.
[14] N. A. Othman et al., "A Hybrid Deep Learning Framework with Decision-Level Fusion for Medical Classification Problems," 2023.
[15] "Transfer-Learning for Herbal/Leaf Classification: Identification of Indian Herbs Using Step-Wise Transfer-Learning," SSRN, 2022.
[16] Biswas et al., "A Comprehensive Review of Explainable AI for Disease-Related Applications," 2024.
[17] Lee et al., "Mobile-Based Medicinal Plant Recognition Using Lightweight CNN Models," 2021.
[18] Patel and S. Banerjee, "Knowledge Graph-Based Symptom–Disease Mapping for AI Healthcare Assistants," 2024.
[19] R. Aravind et al., "Medical Image Data Augmentation: Techniques and Best Practices," 2023.
[20] Mulugeta, A. K. (2024). Deep Learning for Medicinal Plant Species Classification: A Systematic Review.