



A Decision Tree-Based Approach for Disease Prediction and Ayurvedic Drug Recommendation

Suvendu Kumar Nayak¹, Mamata Garanayak², Sangram Keshari Swain³, Gyana Ranjana Panigrahi⁴,
Prabira Kumar Sethy⁵, Aziz Nanthaamornphong^{6*}

¹ Faculty of Engineering and Technology, Sri Sri University, Cuttack, Odisha 754006, India

² Department of Computer Science, Kalinga Institute of Social Sciences, Deemed to be University, Odisha 751024, India

³ Department of Computer Science Engineering, Centurion University of Technology and Management, Odisha 752050, India

⁴ Symbiosis Institute of Technology, Nagpur Campus, Symbiosis International (Deemed University), Pune 440008, India

⁵ Department of Electronics, Sambalpur University, Odisha 768019, India

⁶ College of Computing, Prince of Songkla University, Phuket 83120, Thailand

Corresponding Author Email: aziz.n@phuket.psu.ac.th

Copyright: ©2025 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/isi.300806>

ABSTRACT

Received: 1 June 2025

Revised: 8 August 2025

Accepted: 16 August 2025

Available online: 31 August 2025

Keywords:

Ayurvedic drug recommendation system, decision tree, machine learning, healthcare, personalized medicine

An Ayurvedic drug recommendation system plays a vital role in modern healthcare by offering personalized, holistic treatments based on an individual's unique constitution (Prakriti) and current health imbalances (Vikriti). By integrating ancient Ayurvedic wisdom with contemporary technology, it enhances the accessibility and accuracy of traditional treatments, ensuring they are tailored to each person's lifestyle, diet, and environment. This approach not only predicts disease but also recommends ayurvedic drug for that disease. In our work, we first forecast the sort of ailment that the patient is suffering from using the prediction interface. In the prediction interface, the patient types in his or her present symptoms. The prediction interface then predicts the illness based on the symptoms, and our decision tree model achieves 97% prediction accuracy, 97% of precision and 97% of recall. After determining the disease, ayurvedic medicine is recommended depending on age, gender, disease type, and severity. In our investigation, the recommendation model had a training accuracy of 98%, precision 93%, recall 77% and a testing accuracy of 97%, precision 88% and recall 88%.

1. INTRODUCTION

A recommender system is a technology that suggests items like products, movies, music and books to users based on their preferences. The recommender system is full-fledged systems that takes user data, analyzes it, find patterns and connections and generates recommendations. These are widely used by e-commerce platforms like Amazon or Netflix and many other applications. The primary goal of a recommender system is to suggest relevant and personalized items to users, based on their past behavior, preferences, and similarities with other users [1].

Recommender prototype offers a number of advantages, including better user engagement, time and effort savings, higher discovery of new products, enhanced decision-making for end users, and an overall improvement in their platform or service experience. On the other hands the service provider of recommender system also gets several benefits like cross-selling and upselling opportunities, optimized inventory management, better understanding of user preferences, improved customer retention, increased sales and revenue, targeted marketing and advertising, drive business growth and give them a competitive edge in the market [2].

From individualized treatment recommendations to

operational optimization and nutrition suggestions to hospital recommendations, recommender systems are essential to improving many facets of healthcare delivery and, in the end, improving patient outcomes and experiences [3]. Accurate disease prediction and effective drug recommendation are crucial aspects of preventative and personalized medicine. However, current methods face limitations. According to the World Health Organization (WHO), chronic diseases account for approximately 60% of all deaths globally. Inaccurate or delayed disease diagnosis contributes to treatment inefficacy and increased healthcare costs. Adverse drug reactions (ADRs) are responsible for a significant portion of hospital admissions and patient morbidity. The Institute of Medicine estimates that ADRs contribute to over 100,000 deaths annually in the United States alone. Identifying potentially harmful interactions between medications (Drug - Drug interactions) is also a major challenge. Recommender systems, leveraging machine learning and data analysis, offer promising solutions to disease prediction and drug recommendation and reducing ADRs and DDIs.

Beyond its original borders, Ayurveda's ideas and practices have been disseminated through increased awareness campaigns, educational programs, and wellness tourism efforts. The world over, Ayurveda is becoming more and more

popular as more people become aware of its advantages [4, 5]. Globalization and e-commerce innovations have made Ayurvedic items and practices more widely available. Regardless of their location, consumers may now readily acquire herbal medicines, Ayurvedic supplements, and health services. Since Ayurveda emphasizes natural ingredients and holistic wellbeing, a lot of people are drawn to it because they think it's safer and gentler than mainstream medicine. Ayurvedic treatments are frequently viewed as complimentary or alternative methods of treating long-term illnesses. People believe that Ayurvedic treatments have fewer adverse effects than conventional pharmaceuticals.

While there's growing popularity on Ayurveda, there are several factors contributing to the limited research in developing recommender systems for Ayurvedic medicine. Ayurveda has a more individualized approach based on patient Prakriti (constitution) and Vikriti (imbalance). This makes data collection and standardization for recommender systems difficult [6]. Many Ayurvedic remedies lack the extensive clinical trials required to establish robust scientific evidence for their efficacy. This can make it challenging to develop algorithms that accurately predict treatment outcomes. Ayurvedic diagnosis and treatment consider a wider range of factors than conventional medicine, including dietary habits, lifestyle, and mental state. Capturing these complexities in a way suitable for recommender systems is a challenge.

Effective development of Ayurvedic recommender systems requires collaboration between Ayurvedic practitioners, data scientists, and computer scientists. Developments in artificial intelligence, particularly natural language processing, could facilitate the integration of complex Ayurvedic knowledge into recommender systems.

2. RELATED WORK

Indoriya and Barde [7] develop a system that collect information on 200 herbs and their potential medicinal uses, including relevance for particular diseases. They apply different machine learning (ML) classification methods, such as Support Vector Machines (SVMs), K-Nearest Neighbours (KNN), and Naive Bayes, using the WEKA toolbox. With SVMs exhibiting the greatest performance, the system predicts pertinent herbs for particular ailments with encouraging accuracy. This ML-based method has the potential to be more scalable and widely accessible than the conventional Nadi Pariksha method.

The work of Pogadadanda et al. [8] is a major advancement in the modernization of Ayurvedic medicine. Pulse signal characteristics including frequency, amplitude, and rhythm are extracted by the system. This engine makes use of machine learning techniques that were trained on a collection of clinical and Ayurvedic text data. The engine suggests appropriate Ayurvedic therapies, such as herbal remedies, dietary changes, and lifestyle adjustments, based on the extracted pulse characteristics and patient complaints. Based on pulse analysis, the device shows potential accuracy in diagnosing illnesses.

Majhi et al. [9] introduced a unique method of predicting Parkinson's disease (PD) based on machine learning (ML) and Ayurvedic doshas (Vata, Pitta, and Kapha). The Fox Insight dataset, which includes PD patients' clinical data, was used by the authors. Based on the material that was accessible, Ayurvedic dosha scores were obtained from the MDS-UPDRS-II and MDS-NMSQ scales. Alongside an ensemble

approach, a number of machine learning algorithms (ANN, KNN, SVM, Naive Bayes, Decision Tree, LR and XGBoost) were used. LR, Kernel-SVM, SVM, XGBoost algorithms attained great accuracy (> 92.5%). The model that performed the best was Logistic Regression (LR), which had an accuracy of 92.6% and a reduced false positive rate of 0.045.

Within the context of Ayurvedic principles, Sharoni Narang et al. [10] investigate the potential of machine learning approaches in illness detection. The research fills in the gap between the data-driven methodology of machine learning algorithms and the conventional Ayurvedic diagnostic techniques. It presents a revolutionary framework for the prediction and classification of illnesses based on an individual's constitutional features, symptoms, and other Ayurvedic indicators. Machine learning models are trained on Ayurvedic principles and diagnostic criteria.

Nayak et al. [11] developed a drug recommendation system prototype that used user-entered symptoms to propose potential adverse effects for a given prescription. Four distinct prototypes are used to forecast the diseases. The Vader tool and NLP-based sentiment analysis are used to examine the reviews. Finally, the medications are suggested using probabilistic and weighted average techniques.

In 2021, Manjula and AnandaRaj [12] investigated how machine learning (ML) methods may be applied to Ayurvedic diagnostics. This study sheds insight on the current investigation into the potential integration of artificial intelligence (AI) and machine learning with conventional medicinal systems such as Ayurveda. The goal of the authors' machine learning methods is to automate and optimize the diagnostic process, which might result in recommendations for more effective and precise illness diagnosis and treatment. The accessibility, efficacy, and precision of diagnoses might all be enhanced by this combination.

In order to promote integrative healthcare methods, Madaan and Goyal [13] set out to create machine learning models that could forecast the body's Ayurvedic constituent balance. The 405 people in the dataset that the authors used had tridoshic imbalance diagnoses. Information on symptoms, demographics, and maybe conventional Ayurvedic assessments were included in this data. Numerous conventional machines learning methods, including Support Vector Machine (SVM), K-Nearest Neighbours (KNN), Naive Bayes (NB), Decision Tree (DT), and Artificial Neural Network (ANN), were used in the study. These algorithms were also combined in an ensemble approach that was put into practice. Standard criteria including as accuracy, precision, recall, and F-score were used to assess each model's performance. With a 95% accuracy rate, the ensemble approach performed the best. This shows that, in this particular situation, integrating many machine learning algorithms may result in predictions that are more accurate than using just one method.

Wedaduru is a web-based tool that combines machine learning algorithms and the expertise of Ayurvedic practitioners, created by the Bandara et al. [14]. The Wedaduru system was equipped with machine learning algorithms, natural language processing techniques, and knowledge representation methodologies to facilitate intelligent decision-making. Heart disorders are categorised into four groups, Vataja, Pitaja, Kaphaja, and Krimija, and are predicted with an accuracy rate of over 86%.

Gomes et al. [15] study suggests a revolutionary strategy that integrates machine learning methods with Ayurvedic

medicine from Sri Lanka. To be more precise, it makes use of the deep learning models Inception, InceptionV3, and VGG16 to increase the precision of skin type, skin condition, and severity categorization. The InceptionV3 model achieves an astounding 97% accuracy in identifying different skin conditions based on the results of these tests, while the Inception Resnet model achieves an 86% accuracy in identifying skin types. In addition, the VGG16 model has an amazing 96% accuracy rate for determining the severity of these illnesses. Based on the tenets of ayurveda medicine, this study used a random forest algorithm with a 94.25% accuracy rate to suggest additional therapies.

In order to provide individualised health promotion, Patil and Gore [16] created a unique recommendation method that incorporates yoga and raga therapies specific to each person's constitution. In accordance with the person's physical and mental requirements, these suggestions include certain yoga poses, breathing techniques, food plans, and raga compositions. For creating individualised suggestions, Decision Tree, Naive Bayes, and K-star algorithms are employed 87.60%.

A recommendation system that uses deep learning architectures, including stacked RBMs (Restricted Boltzmann machines), to analyze the collected data and forecast individualized suggestions for physical activity was presented by Bhimavarapu et al. [17]. The experiment includes 658 patients of respiratory illnesses. The suggested recommendation model performs better than the current models, which include the content-based, random forest, SVM, decision tree, and KNN models. Its RMSE and MAE values are 0.0251 and 0.0010, respectively.

The purpose of Hewa et al. [18] is to inform evidence-based strategies for disease prevention and control by offering insightful information on the spread of renal disorders in Sri Lanka. Machine learning approaches, including Random Forest, SVMs, Decision Trees, and others, were applied to 400 patient samples suffering from chronic kidney disease (CKD). There are four distinct components to the research: the first part involves forecasting future hospital medication consumption, diagnosing CKD and suggesting Ayurvedic remedies; the second part involves predicting patient food plans; the third part uses patient symptom information to predict CKD.

A Hybrid Extreme Learning Machine (HELM) was utilised by Chelluboina and Rao [19] to categorise the tongue pictures into several illness groups. By combining elements of conventional and stochastic Extreme Learning Machines, HELM may increase classification accuracy. Grey Level Co-occurrence Matrix (GLCM) and colour moments were used to extract texture and colour attributes. CATDSNet, the suggested model, was assessed for performance using measures such as F1-score, accuracy, precision, and recall. CATDSNet outperforms other conventional machine learning methods in terms of classification accuracy, with 92.3% for diseases such as pancreatitis, heart disease, appendicitis, bronchitis, and gastritis.

An end-to-end framework with several components is proposed by Rathi et al [20]. These components include feature selection, data pre-processing, illness prediction, medication advice, and hospital referral. For illness prediction based on patient data, the methodology makes use of a variety of machine learning techniques, including decision trees, support vector machines, and neural networks. The framework predicts a wide range of illnesses with excellent accuracy

rates, which helps medical professionals identify and treat patients early.

To take use of the network structure for intricate interactions between medications and illnesses, the Yu et al. [21] put up a brand-new LAGCN model. In order to acquire informative representations of medications and illnesses based on their network neighbourhoods, GCN (Graph Convolutional Networks) layers were utilised. Metrics like as AUC (Area Under the ROC Curve), recall, precision, and F1-score were used to evaluate the model. The precision-recall curve area of 0.3168 and the receiver-operating characteristic curve area of 0.8750, respectively, are better achieved by LAGCN than by current cutting-edge prediction techniques and baseline methods.

Rajesh and Rajendran [22] suggested multi-agent reinforcement learning system trained on medical data, enabling the agents to understand the connections between illnesses, symptoms, and efficient therapies. An agent container that can store several mobile agents and gather information from different sources. Different illness classes are created, and the disease attraction weight (DAW) is determined for each class. Recommendations are generated based on the method's prediction of disease and measurement of disease curing rate, or DCR, which is based on DAW.

The intelligent fuzzy inference rule-based predictive diabetes diagnosis (IFIR.PDDM) model developed by Nagaraj and Deepalakshmi [23] focuses on three aspects of diabetes risk prediction. These are the Decision Tree Rule Induction, the Fuzzy Membership Function, and the Fuzzy Inference utilizing Mamdani's Technique. The suggested model is developed and evaluated using the PIMA Indian Diabetes dataset's electronic health record (EHR) medical and clinical information. The method shortens diagnostic times, according to the authors.

The work of Lei et al. [24] suggests ICFCDA, a computational technique to forecast possible relationships between illnesses and circular RNAs. ICFCDA overcomes the drawbacks of conventional techniques by employing a collaborative filtering recommendation system technique. Databases such as circR2Disease provide information on known correlations between circRNA and illness. The accuracy of the predictions made by ICFCDA is evaluated by the authors using methods such as leave-one-out cross-validation. ICFCDA outperforms other current techniques, with an area under the curve of 0.946.

A multi-level decision-making framework for forecasting and suggesting activities linked to heart disease is put out by Sharma and Samant [25]. The heart disease classification system uses the Support Vector Machine algorithm. The accuracy of the proposed MSVM model was 94.09%.

Yu et al. [26] suggest using data gathered through the Internet of Medical Things (IoMT) to create an enhanced DeepFM (Deep Factorization Machine) for illness prediction. In order to concentrate on important elements in the high-dimensional IoMT data, the method includes an attention mechanism. DeepFM design effectively models both low-order and high-order feature interactions by fusing factorization machines with deep neural networks. Sensor readings, ambient conditions, and patient information are probably included in the data. Area under the receiver operating characteristic curve (AUC-ROC) is used to assess the performance of the model. When compared to baseline approaches, the improved model provides greater prediction accuracy.

3. PROPOSED METHODOLOGY

The proposed methodology comprises of two parts. One is disease prediction part and then the Ayurvedic drug recommendation part as shown in Figure 1.

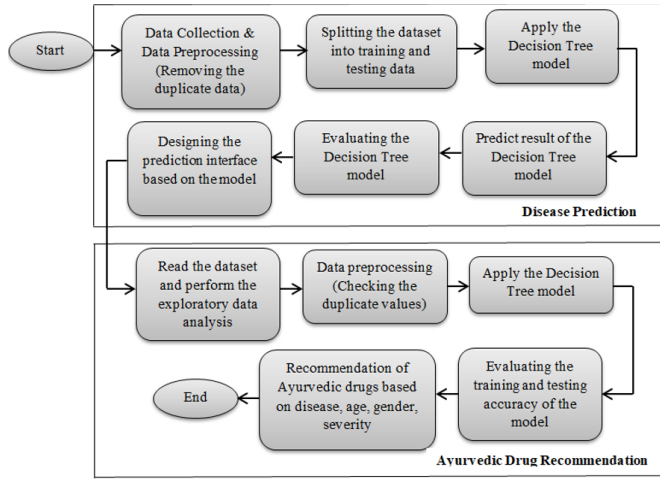


Figure 1. Disease predictions and Ayurvedic drug recommendation

3.1 Proposed disease prediction

Algorithm

Step 1: Load the collected dataset

1. Input Data $M = \{S, P\}$

$S = \{s_1, s_2, s_3, \dots, s_n\}$ are the symptoms vectors (features)

$P = \{p_1, p_2, p_3, \dots, p_n\}$ are the diseases labels (targets)

Step 2: Data pre-processing

2. Data pre-processing

Feature matrix $S \in \mathbb{R}^{n \times m}$, where n is the no. of instances and m is the no. of features.

Target vector $P \in \mathbb{R}^n$.

3. Missing value handling

For all $s_i \in S$, if $s_i = \emptyset$ then:

$s_i \leftarrow \text{mean}(s_i)$

4. If categorical value present: encode

For all $s_i \in S$ (categorical) transform s_i to numerical value

One hot encoding

Step 3: Divide the dataset

5. Split M into training and testing sets:

Training data: $M_r^t = \{S_r^t, P_r^t\}$,

where $|M_r^t| = 0.75 \times |M|$

Testing data: $M_{\text{test}} = \{S_{\text{test}}, P_{\text{test}}\}$,

where $|M_{\text{test}}| = 0.2 \times |M|$

Step 4: Apply decision tree

6. DT classifier T is defined, parameterized by:

(1) Choose Splitting criterion $H \in \text{Entropy}$

where entropy (H) = $\sum_k P_k \log_2(P_k)$

(2) Stopping criteria $\text{max_depth} = 7$

Step 5: Model training

7. To minimize the impurity, optimize the splits

For all nodes in T :

Split nodes relied on feature F that minimized the impurity

I:

$I = I(\text{parent}) - [W_1 \times I(\text{child_node}_1) + W_2 \times I(\text{child_node}_2)]$,

where W_1 and W_2 are the proportions of samples of child nodes.

Step 6: Model evaluation

8. Prediction of class labels for testing test:

For all $s_i \in S_{\text{test}}$:

$\hat{p}_i \leftarrow T.\text{predict}(s_i)$

9. Computation of evaluation metrics:

Accuracy $A: (\sum_i |(\hat{p}_i = p_i)|) / |S_{\text{test}}|$

Evaluate precision, recall and also F_1 score

Step 7: Prediction function

10. Define “predict_disease(symptoms):

Input: Feature vector $S \in \mathbb{R}^m$ (symptoms)

Output: Predicted disease $\hat{p} \in p$

11. Utilize the train tree T to predict:

(1) $\hat{p} = T.\text{predict}(s)$

(2) return $\hat{p}$

Step 8: Create the patient interface

12. Creation of interactive patient's interface

(1) Prompt the patient for symptoms input $s_u \in \mathbb{R}^m$

(2) Normalize and encode s_u

13. Call predict_disease with s_u :

$\hat{p}_u = \text{predict_disease}(s_u)$

Step 9: Display prediction:

Output: predicted disease $\hat{p}_{\text{new}}$

Step 10: Test and validate the interface

14. Test the interface with sample inputs

(1) Utilize known data from M_{test} to validate prediction

(2) Compare $\hat{p}_{\text{test}}$ with p for accuracy

3.2 Data collection, analysis and data preprocessing

Collection of data is the techniques and concepts in preparing raw data for utilization in machine learning and deep learning prototypes. The data can pool by capturing, loading, and web scarping. Large amount of data creation or collection can be the hardest part of machine learning prototypes. For disease prediction we have collected symptoms data (data.csv) from <http://www.kaggle.com/>. The symptoms dataset consists of 32 types of symptoms like acidity, headache, excessive hunger, muscle weakness etc. of 1071 patients as shown in Figure 2.

	acidity	indigestion	headache	blurred_and_distorted_vision	excessive_hunger	muscle_weakness	stiff_neck	swelling_joints	movement_stiffness	depress
0	1	1	0	1	1	0	1	0	0	0
1	0	1	1	1	1	0	1	0	0	0
2	1	0	1	1	1	0	1	0	0	0
3	1	1	0	1	1	0	1	0	0	0
4	1	1	1	0	1	1	1	0	0	0
...	...	...	...	...	...	...	...	...	...	...
1066	0	1	0	0	0	0	0	0	0	0
1067	1	1	0	0	0	0	0	0	0	0
1068	0	1	0	0	0	0	0	0	0	0
1069	0	1	0	0	0	0	0	0	0	0
1070	0	1	0	0	0	0	0	0	0	0
1071	0	1	0	0	0	0	0	0	0	0

1071 rows x 32 columns

```
print(df1.columns.tolist())
```

['acidity', 'indigestion', 'headache', 'blurred_and_distorted_vision', 'excessive_hunger', 'muscle_weakness', 'stiff_neck', 'swelling_joints', 'movement_stiffness', 'depression', 'irritability', 'visual_disturbances', 'painful_walking', 'abdominal_pain', 'nausea', 'vomiting', 'blood_in_mucus', 'fatigue', 'fever', 'dehydration', 'loss_of_appetite', 'cramping', 'blood_in_stool', 'g_nawing', 'upper_abdominal_pain', 'fullness_feeling', 'hiccups', 'abdominal_bloating', 'heartburn', 'belching', 'burning_ache', 'prognosis']

Figure 2. Symptoms data

Duplicate Rows :

	acidity	indigestion	headache	blurred_and_distorted_vision	excessive_hunger	muscle_weakness	stiff_neck	swelling_joints	movement_stiffness	depress
22	1	0	1	1	1	0	1	0	0	0
39	1	1	1	1	0	0	1	0	0	0
84	1	0	1	1	0	0	1	0	0	0
85	1	1	1	1	1	1	0	0	0	0
93	1	1	1	1	0	0	1	0	0	0
...	...	...	...	...	...	...	...	...	...	...
1053	1	0	0	0	0	0	0	0	0	0
1054	1	1	0	0	0	0	0	0	0	0
1058	0	1	0	0	0	0	0	0	0	0
1064	1	1	0	0	0	0	0	0	0	0
1068	0	1	0	0	0	0	0	0	0	0

295 rows x 32 columns

Figure 3. Duplicate data

Data preprocessing increases the accuracy and quality of data by removing computer mistakes and missing or inconsistent information that was gathered by humans. Additionally, it produces the reliable data. It improves accuracy and reliability. For the preprocessing we first check the duplicate data and as the result we found 295 duplicate data as in Figure 3.

After detection of the duplicate data, we simply remove the duplicate data from the original dataset. Finally, after the removal of data we have 776 data in the dataset as in Figure 4.

	acidity	indigestion	headache	blurred_and_distorted_vision	excessive_hunger	muscle_weakness	stiff_neck	swelling_joints	movement_stiffness	depress
0	1	1	0	1	1	0	1	0	0	0
1	0	1	1	1	1	0	1	0	0	0
2	1	0	1	1	1	0	1	0	0	0
3	1	1	0	1	1	0	1	0	0	0
4	1	1	1	0	1	1	1	0	0	0
...	...	...	...	...	...	...	...	...	...	...
1065	0	1	0	0	0	0	0	0	0	0
1066	0	1	0	0	0	0	0	0	0	0
1067	1	1	0	0	0	0	0	0	0	0
1068	0	1	0	0	0	0	0	0	0	0
1070	0	1	0	0	0	0	0	0	0	0

776 rows x 32 columns

Figure 4. Final dataset

Finally, the data set was checked for the null values and there were no null values present in the dataset. Figure 5 shows the feature correlation of symptoms.

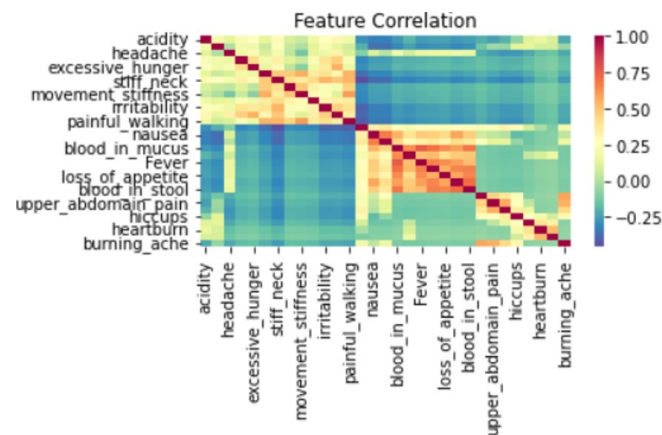


Figure 5. Feature correlation of symptoms

After the data is pre-processed, the processed DataFrame is converted into a NumPy array, splitting the features of the dataset as well as the labels of the dataset and finally splitting features and labels as training and testing data.

3.3 Apply the decision tree model for disease prediction

```
drugTree = DecisionTreeClassifier(criterion="entropy", max_depth = 7)
drugTree

DecisionTreeClassifier
DecisionTreeClassifier(criterion='entropy', max_depth=7)

fitting the data with the training feature matrix X_trainset and training response vector y_trainset

drugTree.fit(X_trainset,y_trainset)

DecisionTreeClassifier
DecisionTreeClassifier(criterion='entropy', max_depth=7)
```

Figure 6. Decision tree model

A flexible machine learning approach that may be applied to both the regression and classification problems is the decision tree. It's among the most straightforward and natural

ways to use predictive modelling. In our method, we have used the decision tree classifier with criterion = 'entropy' and max_depth = 7 for prediction and then fit the training dataset and testing dataset to the decision tree model as in Figure 6.

3.4 Prediction outcome

After the data is fitted to the decision tree model, the decision tree model predicted the diseases as the outcome of the model as in Figure 7.

```
actual target : ['Gastritis' 'Arthritis' 'Arthritis' 'Migraine' 'Arthritis' 'Gastritis'
'Gastritis' 'Migraine' 'Gastritis' 'Gastritis' 'Arthritis' 'Diarrhea'
'Diarrhea' 'Arthritis' 'Gastritis' 'Arthritis' 'Migraine' 'Arthritis'
'Arthritis' 'Arthritis' 'Migraine' 'Arthritis' 'Arthritis' 'Arthritis'
'Migraine' 'Diarrhea' 'Gastritis' 'Gastritis' 'Migraine' 'Migraine'
'Gastritis' 'Gastritis' 'Gastritis' 'Migraine' 'Gastritis' 'Arthritis'
'Migraine' 'Gastritis' 'Migraine' 'Arthritis' 'Diarrhea' 'Arthritis'
'Diarrhea' 'Diarrhea' 'Diarrhea' 'Gastritis' 'Gastritis' 'Gastritis'
'Gastritis' 'Arthritis' 'Diarrhea' 'Migraine' 'Diarrhea' 'Arthritis'
'Diarrhea' 'Diarrhea' 'Gastritis' 'Migraine' 'Diarrhea' 'Arthritis'
'Migraine' 'Gastritis' 'Arthritis' 'Migraine' 'Arthritis' 'Gastritis'
'Arthritis' 'Arthritis' 'Arthritis' 'Gastritis' 'Gastritis' 'Migraine'
'Gastritis' 'Arthritis' 'Migraine' 'Diarrhea' 'Migraine' 'Arthritis'
'Gastritis' 'Gastritis' 'Gastritis' 'Gastritis' 'Diarrhea' 'Gastritis'
'Gastritis' 'Migraine' 'Arthritis' 'Diarrhea' 'Gastritis' 'Migraine'
'Gastritis' 'Diarrhea' 'Arthritis' 'Migraine' 'Gastritis' 'Migraine'
'Arthritis' 'Diarrhea' 'Arthritis' 'Arthritis' 'Migraine' 'Gastritis'
'Gastritis' 'Gastritis' 'Arthritis' 'Gastritis' 'Migraine' 'Migraine'
'Gastritis' 'Diarrhea' 'Arthritis' 'Arthritis' 'Migraine' 'Diarrhea'
'Migraine' 'Migraine' 'Migraine' 'Migraine' 'Diarrhea' 'Migraine'
'Gastritis' 'Diarrhea' 'Diarrhea' 'Migraine' 'Arthritis' 'Migraine'
'Migraine' 'Gastritis' 'Migraine' 'Diarrhea' 'Migraine' 'Diarrhea'
'Diarrhea' 'Diarrhea' 'Arthritis' 'Migraine' 'Migraine' 'Migraine'
'Gastritis' 'Migraine' 'Diarrhea' 'Migraine' 'Arthritis' 'Migraine'
'Arthritis' 'Gastritis' 'Diarrhea' 'Diarrhea' 'Gastritis' 'Arthritis'
'Arthritis' 'Gastritis' 'Diarrhea' 'Diarrhea' 'Gastritis' 'Arthritis'
'Gastritis' 'Migraine' 'Migraine' 'Arthritis' 'Diarrhea' 'Gastritis'
'Migraine' 'Migraine' 'Migraine' 'Migraine' 'Diarrhea' 'Gastritis'
'Arthritis' 'Gastritis' 'Arthritis' 'Migraine' 'Gastritis' 'Diarrhea'
'Gastritis' 'Arthritis' 'Arthritis' 'Migraine' 'Migraine' 'Diarrhea'
'Migraine' 'Arthritis' 'Migraine' 'Diarrhea' 'Gastritis' 'Migraine'
'Gastritis' 'Gastritis' 'Gastritis' 'Migraine' 'Gastritis' 'Arthritis'
'Migraine' 'Gastritis' 'Migraine' 'Arthritis' 'Arthritis' 'Migraine'
'Diarrhea' 'Diarrhea' 'Gastritis' 'Migraine' 'Arthritis' 'Gastritis'
'Gastritis' 'Gastritis' 'Gastritis' 'Migraine' 'Diarrhea' 'Gastritis'
'Gastritis' 'Migraine' 'Arthritis' 'Arthritis' 'Gastritis' 'Migraine'
'Arthritis' 'Arthritis' 'Arthritis' 'Gastritis' 'Arthritis' 'Migraine'
'Gastritis' 'Arthritis' 'Migraine' 'Diarrhea' 'Migraine' 'Arthritis'
'Gastritis' 'Diarrhea' 'Gastritis' 'Migraine' 'Gastritis' 'Diarrhea'
'Arthritis' 'Gastritis' 'Arthritis' 'Gastritis' 'Migraine' 'Migraine'
'Migraine' 'Migraine' 'Diarrhea' 'Migraine' 'Gastritis' 'Gastritis'
'Arthritis' 'Migraine' 'Arthritis' 'Diarrhea' 'Migraine' 'Diarrhea'
'Arthritis' 'Gastritis' 'Diarrhea' 'Migraine' 'Arthritis' 'Migraine'
'Gastritis' 'Migraine' 'Diarrhea' 'Migraine' 'Arthritis' 'Migraine'
'Arthritis' 'Gastritis' 'Diarrhea' 'Migraine' 'Gastritis' 'Migraine'
'Gastritis' 'Gastritis' 'Gastritis' 'Migraine' 'Gastritis' 'Diarrhea'
'Arthritis' 'Gastritis' 'Arthritis' 'Migraine' 'Migraine' 'Diarrhea'
'Migraine' 'Arthritis' 'Migraine' 'Diarrhea' 'Arthritis']
```

Figure 7. Prediction of disease

3.5 Evaluation of prediction model

It refers to the process of assessing the performance of our proposed model on a separate, unseen dataset, typically called the testing data. This evaluation helps to determine how well our model generalizes to new data, which is crucial for ensuring its effectiveness and reliability in real-world applications. In this proposed work, the classification report is shown in Figure 8, which shows overall 97% accuracy. The

classification report also shows the precision 92%, recall 98%, F1-score 95% for Arthritis, precision 98%, recall 100%, F1-score 99% for Diarrhea, precision 100%, recall 98%, F1-score 99% for Gastritis, precision 98%, recall 92%, F1-score 95% for Migraine.

Classification Report :

	precision	recall	f1-score	support
Arthritis	0.92	0.98	0.95	58
Diarrhea	0.98	1.00	0.99	48
Gastritis	1.00	0.98	0.99	64
Migraine	0.98	0.92	0.95	63
accuracy			0.97	233
macro avg	0.97	0.97	0.97	233
weighted avg	0.97	0.97	0.97	233

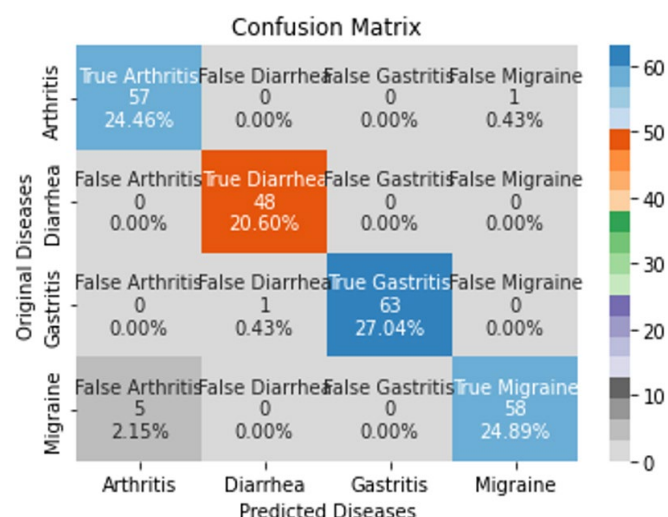


Figure 8. Classification report and confusion matrix

4. PREDICTION INTERFACE

For making an interface we first create a function `predict_disease_from_symptom ()` which takes a list of symptoms as input and predicts the corresponding disease using a pre-trained Decision Tree model. It first initializes a dictionary where all possible symptoms are set to zero, representing absence. For each symptom present in the input list, the corresponding dictionary value is updated to one, indicating presence. This dictionary is then converted into a single-row DataFrame, where each column corresponds to a symptom and the row contains binary values representing whether that symptom is present or not. The function then loads a pre-trained Decision Tree model from a saved file and utilizes it to predict the disease based on the input symptom data. After making the prediction, the temporary test DataFrame is deleted to free up memory, and finally, the predicted disease label is returned as output.

Then an inference is created using Gradio to create a simple web-based interface for disease prediction. The interface connects to the previously defined function `predict_disease_from_symptom ()`, which takes user-selected symptoms as input and returns the predicted disease. The input is provided through a CheckboxGroup, where users can select one or more symptoms from a predefined list, and the output is displayed as plain text. A short description is added to guide

users, instructing them to select symptoms from the list to receive a prediction. Finally, `iface.launch(share = True)` starts the server and generates a publicly accessible URL, allowing the demo to be run in a separate page or shared with others for testing and validation.

The disease prediction interface is shown in Figure 9 and Figure 10 where the patients have to enter his/her symptoms. More than one symptom can also be added in the interface. Based on the symptoms of the patient, the interface predicts the disease from which the patient is suffering for.

Figure 9. Prediction interface predicting disease diarrhea

Figure 10. Prediction interface predicting disease arthritis

5. PROPOSED AYURVEDIC DRUG RECOMMENDATION

5.1 Dataset and exploratory data analysis

Because of the traditional and holistic character of Ayurvedic therapy, there are various specific issues while collecting statistics for Ayurvedic studies. Ayurveda, a medical system having historical roots in the Indian subcontinent, focuses on natural therapies and overall well-being. The data we have collected from [http://www.kaggle.com// Drug prescription Dataset.csv](http://www.kaggle.com//Drug%20prescription%20Dataset.csv) which consists of 1288 rows having different diseases and 5 columns having disease, age, gender, severity and Ayurvedic drug as shown in Figure 11.

The dataset consists of one attribute known as severity. The severity column consists of 3 values; high, low severity and normal as in Figure 12.

Figure 13 shows the total number of male and female patients.

Figure 14 shows the categories and the total number of diseases. There are totally four types of diseases; arthritis (334), diarrhea (318), gastritis (318), migraine (318). Different Types of Ayurvedic Drugs are shown in Figure 15.

	disease	age	gender	severity	drug
0	diarrhea	4	male	LOW	promegranate drink
1	diarrhea	4	male	NORMAL	lime juice
2	diarrhea	5	male	LOW	promegranate drink
3	diarrhea	5	male	NORMAL	lime juice
4	diarrhea	6	male	LOW	promegranate drink
...	...	...	...	...	...
1283	diarrhea	60	male	NORMAL	kutajarishta
1284	diarrhea	60	male	LOW	ajamodarka
1285	diarrhea	60	female	HIGH	sitopaladi churna
1286	diarrhea	60	female	NORMAL	kutajarishta
1287	diarrhea	60	female	LOW	ajamodarka

1288 rows × 5 columns

Figure 11. Ayurvedic drug dataset

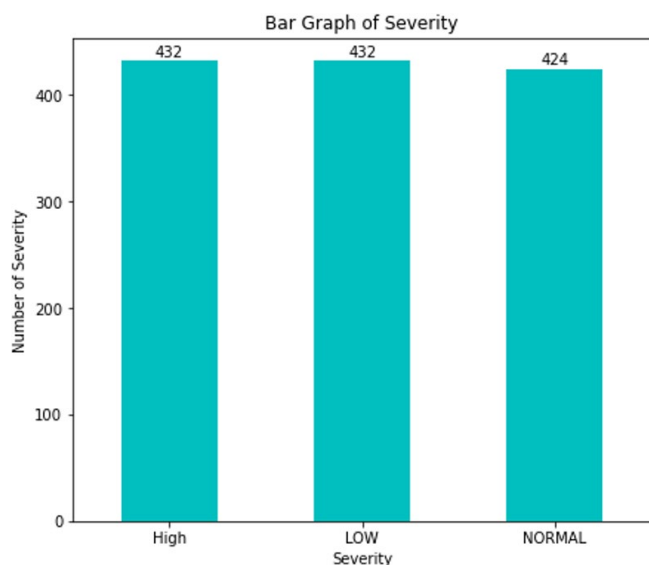


Figure 12. Severity graph

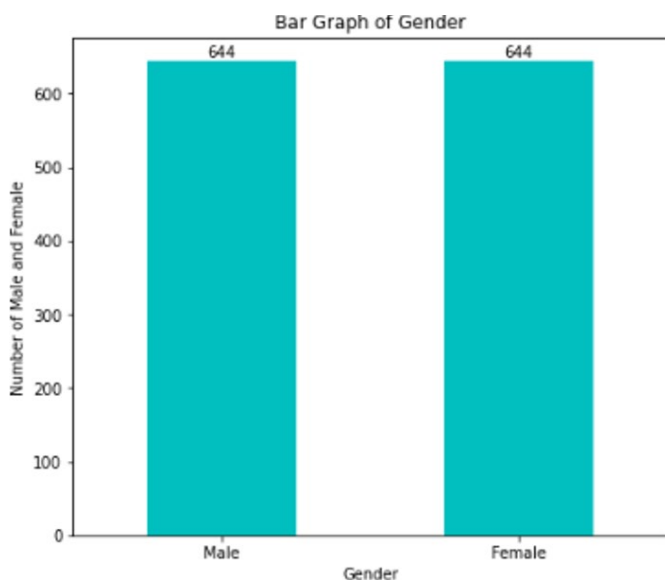


Figure 13. Number of male and female patients

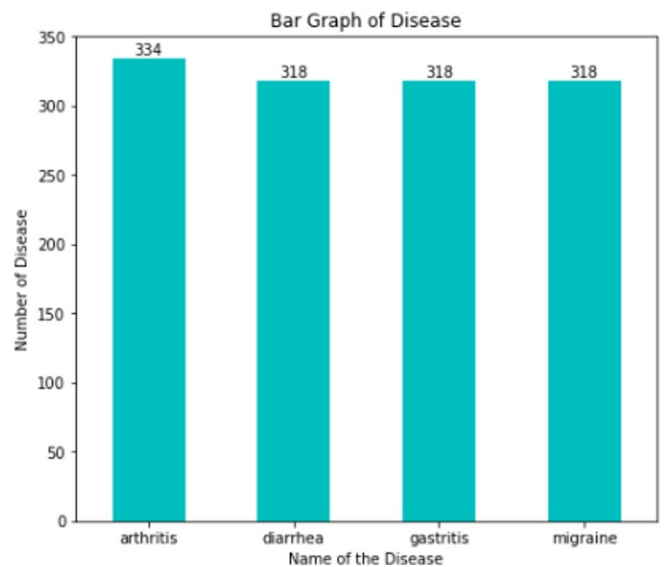


Figure 14. Types of diseases

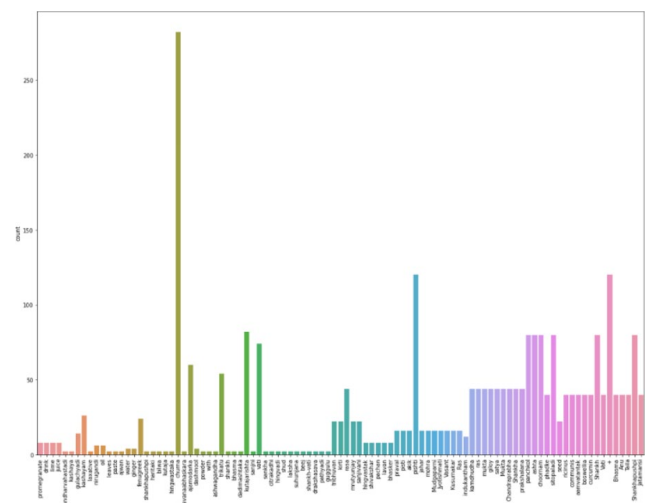


Figure 15. Different types of ayurvedic drugs

5.2 Data preprocessing

Preparation of data entails converting raw data to a clean and useful format. In order to assure the accuracy and efficiency of our analytical procedures we first checked the presence of duplicate values and removed them. After that we removed the column containing the target names such as "disease", "age", "gender", "severity" since it doesn't contain numeric values. Then converting categorical variables such as "disease", "gender", "severity" into dummy/indicator variables. The conversion details are given below.

Disease: 0: arthritis, 1: diarrhea, 2: gastritis, 3: migraine

Gender: 0: male, 1: female

Severity: 0: low, 1: medium, 2: high

5.3 Apply the DT model for recommendation

To apply the DT model, we first Setting a random seed = 42 for reproducibility. Finally, we have split the data set into training and testing set in the ratio of test size=0.3 and save the train and test file in different .csv files. Finally, we model the decision tree classifier with criterion = "entropy" and max_depth = 6. The choice of max-depth = 6 is made to strike a balance between model complexity and overfitting.

Shallower trees (depth < 5) exhibited underfitting with lower accuracy (~ 72%), while deeper trees (depth > 8) tended to memorize the training data, leading to a significant performance gap between training and validation accuracy (> 11%). A depth of 6 provided the most stable generalization, with validation accuracy of 97%. The criterion = 'entropy' is chosen over the default Gini index to maximize information gain at each split. The DT model predicts the following prediction as in Figure 16.

```
[ 'panchkol churna' 'Shankh Vati + Shankh Bhasma' 'kutajarishta'
'jahar mohra pishti' 'Mukta pishti' 'sanjivani vati' 'sitopaladi churna'
'phadke' 'mukta pishti' 'giloy satya' 'Shankha prakshalana'
'mukta pishti' 'ajamodarka' 'mukta pishti' 'aamvatantak churna'
'kutajarishta' 'Anu Taila + Shankhapushpi' 'ashta choornam'
'jahar mohra pishti' 'ajamodarka' ]
[ 'panchkol churna' 'Shankh Vati + Shankh Bhasma' 'kutajarishta'
'jahar mohra pishti' 'Mukta pishti' 'sanjivani vati' 'sitopaladi churna'
'phadke' 'mukta pishti' 'giloy satya' 'Shankha prakshalana'
'mukta pishti' 'ajamodarka' 'mukta pishti' 'aamvatantak churna'
'kutajarishta' 'Anu Taila + Shankhapushpi' 'ashta choornam'
'suhunjana beej' 'ajamodarka' ]
```

Figure 16. Prediction of ayurvedic drugs using DT

5.4 Evaluating the training and testing accuracy of the model

The model validation is done through internal validation. The dataset was split into training (70%) and testing (30%) subsets. Evaluation metrics such as Precision, Recall, F1-score were computed.

	precision	recall	f1-score	support
Anu Taila + Shankhapushpi	1.00	1.00	1.00	24
Chandraprabha vati	1.00	1.00	1.00	31
Jatamansi + Shankhapushpi	1.00	1.00	1.00	24
Jyotishmati	1.00	1.00	1.00	11
Mudgaparni	1.00	1.00	1.00	11
Mukta pishti	1.00	1.00	1.00	33
Shankh Vati + Shankh Bhasma	1.00	1.00	1.00	33
Shankha prakshalana	1.00	1.00	1.00	30
Vasant Kusumakar Ras	1.00	1.00	1.00	11
aamvatantak churna	1.00	1.00	1.00	28
ajamodarka	0.90	1.00	0.95	36
ajwan water	0.33	1.00	0.50	1
akik pishti	1.00	1.00	1.00	11
ashta choornam	1.00	1.00	1.00	58
ashwagandha	0.40	1.00	0.57	2
bilwa	1.00	1.00	1.00	1
boswellia curcumin	1.00	1.00	1.00	26
citrakadhi vati	0.50	1.00	0.67	1
dadimashtaka churna	1.00	1.00	1.00	1
dashmool oil	1.00	0.00	0.00	1
dashmool powder with water	1.00	0.00	0.00	1
drakshasava	1.00	0.00	0.00	1
fenugreek	1.00	1.00	1.00	2
fenugreek seed	1.00	1.00	1.00	16
gandharvahastadi kashaya	1.00	0.00	0.00	1
giloy satya	1.00	1.00	1.00	31
ginger	1.00	0.00	0.00	2
ginger	0.40	1.00	0.57	2
gulachyadi kashayam	1.00	1.00	1.00	12
haritaki	1.00	1.00	1.00	1
hingvadi vati	1.00	0.00	0.00	2
hingvastak churna	1.00	1.00	1.00	5
hingvastaka churna	0.40	1.00	0.57	2
indukantham kashayam	1.00	1.00	1.00	6
jahar mohra pishti	1.00	1.00	1.00	11
kamdudha ras	1.00	1.00	1.00	28
kutaja	1.00	0.00	0.00	2
kutajarishta	1.00	1.00	1.00	58
lavan bhasker churna	1.00	1.00	1.00	6
lavanaabhasakara churna	1.00	0.00	0.00	1
laxative	1.00	0.00	0.00	2
lime juice	1.00	1.00	1.00	8
mrutyunjay rasa	1.00	1.00	1.00	17
mukta pishti	1.00	1.00	1.00	30
nirugandi leaves paste	1.00	0.00	0.00	2
nirugandi oil	1.00	1.00	1.00	3
panchkol churna	1.00	1.00	1.00	57
phadke	1.00	1.00	1.00	33
praval pisti	1.00	1.00	1.00	11
promegranate drink	0.57	1.00	0.73	4
ricinus communis	1.00	1.00	1.00	28
sanjivani vati	1.00	1.00	1.00	10
sanjni vati	1.00	0.00	0.00	1
sankha vati	1.00	0.00	0.00	2
shankh bhasma	0.50	1.00	0.67	2

shankh-vati	1.00	0.00	0.00	1
shankhapushpi	1.00	0.00	0.00	2
shivakshar pachan churna	1.00	1.00	1.00	8
shud laksha	1.00	1.00	1.00	2
sitopaladi churna	1.00	1.00	1.00	61
suhunjana beej	1.00	0.00	0.00	1
tribhuvan kirti rasa	1.00	1.00	1.00	11
trikatu	0.67	1.00	0.80	2
trikatu churna	1.00	1.00	1.00	39
accuracy			0.98	901
macro avg	0.93	0.77	0.72	901
weighted avg	0.99	0.98	0.97	901

Figure 17. Training accuracy of DT model

Confusion matrix (training):

```
[[24 0 0 ... 0 0 0]
[ 0 31 0 ... 0 0 0]
[ 0 0 24 ... 0 0 0]
...
[ 0 0 0 ... 11 0 0]
[ 0 0 0 ... 0 2 0]
[ 0 0 0 ... 0 0 39]]
```

Figure 18. Confusion matrix of training data

Figure 17 shows the training accuracy of DT model. Figure 18 shows the confusion matrix of training data. Figure 19 shows the testing accuracies of DT model and Figure 20 shows the confusion matrix of the testing data.

	precision	recall	f1-score	support
Anu Taila + Shankhapushpi	1.00	1.00	1.00	14
Chandraprabha vati	1.00	1.00	1.00	11
Jatamansi + Shankhapushpi	1.00	1.00	1.00	12
Jyotishmati	1.00	1.00	1.00	3
Mudgaparni	1.00	1.00	1.00	5
Mukta pishti	1.00	1.00	1.00	8
Shankh Vati + Shankh Bhasma	1.00	1.00	1.00	6
Shankha prakshalana	1.00	1.00	1.00	13
Vasant Kusumakar Ras	1.00	1.00	1.00	5
aamvatantak churna	1.00	1.00	1.00	11
ajamodarka	1.00	1.00	1.00	20
ajwan water	1.00	1.00	1.00	1
akik pishti	1.00	1.00	1.00	4
ashta choornam	1.00	1.00	1.00	19
ashwagandha	0.00	1.00	0.00	0
bilwa	1.00	1.00	1.00	1
boswellia curcumin	1.00	1.00	1.00	11
citrakadhi vati	1.00	1.00	1.00	1
dadimashtaka churna	1.00	1.00	1.00	1
dashmool oil	0.50	1.00	0.67	1
dashmool powder with water	1.00	0.00	0.00	1
drakshasava	1.00	1.00	1.00	1
fenugreek seed	1.00	1.00	1.00	3
gandharvahastadi kashaya	1.00	0.00	0.00	1
giloy satya	1.00	1.00	1.00	12
ginger	0.00	1.00	0.00	0
gulachyadi kashayam	1.00	1.00	1.00	2
haritaki	1.00	1.00	1.00	1
hingvastak churna	1.00	1.00	1.00	2
hingvastaka churna	0.00	1.00	0.00	0
indukantham kashayam	1.00	1.00	1.00	6
jahar mohra pishti	1.00	1.00	1.00	5
kamdudha ras	1.00	1.00	1.00	12
kutaja	0.00	1.00	0.00	0
kutajarishta	1.00	0.89	0.94	18
lavan bhasker churna	0.33	1.00	0.50	1
lavanaabhasakara churna	1.00	1.00	1.00	1
mrutyunjay rasa	0.60	1.00	0.75	3
mukta pishti	1.00	1.00	1.00	11
nirugandi oil	1.00	1.00	1.00	1
panchkol churna	1.00	1.00	1.00	20
pathyadi guggulu	1.00	0.00	0.00	2
phadke	1.00	1.00	1.00	5
praval pisti	1.00	1.00	1.00	5
promegranate drink	1.00	1.00	1.00	3
ricinus communis	1.00	1.00	1.00	7
sanjivani vati	1.00	1.00	1.00	11
sanjni vati	1.00	0.00	0.00	1
shankh bhasma	0.00	1.00	0.00	0
shankh-vati	1.00	0.00	0.00	1
sitopaladi churna	1.00	1.00	1.00	16
suhunjana beej	1.00	0.00	0.00	1
tribhuvan kirti rasa	1.00	1.00	1.00	10
trikatu churna	1.00	0.83	0.91	12

accuracy			0.97	322
macro avg	0.88	0.88	0.77	322
weighted avg	0.99	0.97	0.97	322

Figure 19. Testing accuracy of DT model

Confusion Matrix:

```
[[14  0  0 ...  0  0  0]
 [ 0 11  0 ...  0  0  0]
 [ 0  0 12 ...  0  0  0]
 ...
 [ 0  0  0 ...  0  0  0]
 [ 0  0  0 ...  0 10  0]
 [ 0  0  0 ...  0  0 10]]
```

Figure 20. Confusion matrix of testing data

The training accuracy of decision tree model is found 98%.
The testing accuracy of decision tree model is found 97%.

5.5 Recommendation of Ayurvedic drugs based on disease, age, gender, severity

Figure 21 shows the final recommendation.

```
print(model.predict([[2, 14, 0, 0]]))
```

```
['hingvastak churna']
```

```
print(model.predict([[3, 5, 1, 0]]))
```

```
['shankh bhasma']
```

```
print(model.predict([[2, 8, 0, 2]]))
```

```
['fenugreek']
```

Figure 21. Confusion matrix of testing data

In first predict part of Figure 21, 2 indicates the disease type, 14 is the age, 0 indicates male person and last 0 indicates low severity. The first recommended ayurvedic drug is hingvastak churna. In the second part of the figure 21, 20, 3 indicates the disease type, 5 is the age, 1 indicates female person and last 0 indicates low severity. The second recommended ayurvedic drug is shankh bhasma. Likewise, the model will recommend the ayurvedic drug based on disease type, age, male/female and severity (low, medium, high).

6. CONCLUSION AND FUTURE WORKS

Ayurvedic medicine is a holistic approach to health that emphasizes balance and wellness through personalized treatments. By focusing on the balance of the body's energies (doshas) and using natural herbs and lifestyle modifications, Ayurveda aims to treat the root cause of diseases and promote

overall health and well-being. In our work, we first predict the type of the disease, the patient suffers in through the prediction interface. In the prediction interface, the patient enters his or her current symptoms. The prediction interface then guesses the disease based on the symptoms, with our decision tree model achieving a prediction accuracy of 97%. After estimating the ailment, the ayurvedic medication is prescribed based on age, gender, disease kind, and severity. In our study, the recommendation model's training accuracy is 98%, and its testing accuracy is 97% which is better as compare to SVM 92.5% in the reference [9], Ensemble 95% in the reference [13].

In future, this model can be enhanced with advanced ML algorithms and AI models to improve personalization of recommendation based on patient's health profiles, genetics and specific health conditions. Incorporating large scale data from historical Ayurvedic texts, natural language processing can be applied to extract important information from ancient text and can also integrate it with modern medical knowledge. Most often Ayurveda considers aspects such as lifestyle, Vikriti and Prakriti in diagnosis, but our model utilizes the features age, gender, disease and severity. We select this simple model due to the challenges in gathering data on Prakriti and lifestyle at this stage. In future, we plan to extend the model by incorporating these features so that it better reflects the traditional Ayurvedic approach. While the current study demonstrates promising results through experimental validation, a critical next step is comprehensive clinical validation.

REFERENCES

- [1] Roy, D., Dutta, M. (2022). A systematic review and research perspective on recommender systems. *Journal of Big Data*, 9(1): 59. <https://doi.org/10.1186/s40537-022-00592-5>
- [2] Lim, Y.F., Haw, S.C., Ng, K.W., Anaam, E.A. (2023). Hybrid-based recommender system for online shopping: A review. *Journal of Engineering Technology and Applied Physics*, 5(1): 12-34.
- [3] Tran, T.N.T., Felfernig, A., Trattner, C., Holzinger, A. (2021). Recommender systems in the healthcare domain: State-of-the-art and research issues. *Journal of Intelligent Information Systems*, 57(1): 171-201. <https://doi.org/10.1007/s10844-020-00633-6>
- [4] Mukherjee, P.K., Harwansh, R.K., Bahadur, S., Banerjee, S., Kar, A., Chanda, J., Biswas, S., Ahmmed, S.M., Katiyar, C.K. (2017). Development of Ayurveda—Tradition to trend. *Journal of ethnopharmacology*, 197: 10-24. <https://doi.org/10.1016/j.jep.2016.09.024>
- [5] Chattopadhyay, K. (2019). Globalisation of Ayurveda: Importance of scientific evidence base. In *Herbal Medicine in India: Indigenous Knowledge, Practice, Innovation and its Value*, pp. 3-7. Singapore: Springer Singapore. https://doi.org/10.1007/978-981-13-7248-3_1
- [6] Chauhan, A., Semwal, D.K., Mishra, S.P., Semwal, R.B. (2015). Ayurvedic research and methodology: Present status and future strategies. *AYU (An International Quarterly Journal of Research in Ayurveda)*, 36(4): 364-369. <https://doi.org/10.4103/0974-8520.190699>
- [7] Indoriya, P., Barde, S. (2022). Prediction of ayurvedic herbs for specific diseases by classification techniques in

- machine learning. In 2022 2nd Asian Conference on Innovation in Technology (ASIANCON), Ravet, India, pp. 1-4. <https://doi.org/10.1109/ASIANCON55314.2022.9908981>
- [8] Pogadadanda, H., Shankar, U.S., Jansi, K.R. (2021). Disease diagnosis using ayurvedic pulse and treatment recommendation engine. In 2021 7th international conference on advanced computing and communication systems (ICACCS), Coimbatore, India, pp. 1254-1258. <https://doi.org/10.1109/ICACCS51430.2021.9441843>
- [9] Majhi, V., Choudhury, B., Saha, G., Paul, S. (2023). Development of a machine learning-based Parkinson's disease prediction system through Ayurvedic dosha analysis. *International Journal of Ayurvedic Medicine*, 14(1): 180-189.
- [10] Sharoni Narang, S.P., Batwal, O., Khandagale, M. (2018). Ayurveda based disease diagnosis using machine learning. *International Research Journal of Engineering and Technology*, 5(3): 3704-3707.
- [11] Nayak, S.K., Garanayak, M., Swain, S.K., Panda, S.K., Godavarthi, D. (2023). An intelligent disease prediction and drug recommendation prototype by using multiple approaches of machine learning algorithms. *IEEE Access*, 11: 99304-99318. <https://doi.org/10.1109/ACCESS.2023.3314332>
- [12] Manjula, H.M., AnandaRaj, S.P. (2021). Ayurvedic diagnosis using machine learning techniques to examine the diseases by extracting the data stored in AyurDataMart. In 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), Greater Noida, India, pp. 239-244. <https://doi.org/10.1109/ICAC3N53548.2021.9725473>
- [13] Madaan, V., Goyal, A. (2020). Predicting ayurveda-based constituent balancing in human body using machine learning methods. *IEEE Access*, 8: 65060-65070. <https://doi.org/10.1109/ACCESS.2020.2985717>
- [14] Bandara, R.I.S., Prabakaran, S., Perera, S.A.K.G., Banu, M.R., Kahandawaarachchi, K.A.D.C.P. (2019). Wedaduru-An intelligent ayurvedic disease screening and remedy analysis solution. In 2019 International Conference on Advancements in Computing (ICAC), Malabe, Sri Lanka, pp. 151-155. <https://doi.org/10.1109/ICAC49085.2019.9103410>
- [15] Gomes, M.P.O.M., Jayasekara, Y.N., Kariyapperuma, K.M.K.R., Gunawardhna, H.P.M.N., Suwarnakantha, N.R.S., Wimalarathne, G. (2023). Human skin diseases identification and treatment suggestion by Sri Lankan ayurveda medicine using machine learning. In 2023 5th International Conference on Advancements in Computing (ICAC), Colombo, Sri Lanka, pp. 65-70. <https://doi.org/10.1109/ICAC60630.2023.10417632>
- [16] Patil, P., Gore, S. (2016). Recommendation system for yoga and raga for personalized health based on constitution. In 2016 International Conference on Computing Communication Control and automation (ICCUBEA), Pune, India, pp. 1-5. <https://doi.org/10.1109/ICCUBEA.2016.7860010>
- [17] Bhimavarapu, U., Sreedevi, M., Chintalapudi, N., Battineni, G. (2022). Physical activity recommendation system based on deep learning to prevent respiratory diseases. *Computers*, 11(10): 150. <https://doi.org/10.3390/computers11100150>
- [18] Hewa, H.K.L., Deshan, N.A.S., Swarnakantha, N.P.R.S., Kumari, S. (2023). Machine learning approach to identify spread of kidney diseases in Sri Lanka. In 2023 5th International Conference on Advancements in Computing (ICAC), Colombo, Sri Lanka, pp. 83-88. <https://doi.org/10.1109/ICAC60630.2023.10417359>
- [19] Chelluboina, S.P., Rao, K.N. (2023). CATDSNet: Computer aided tongue diagnosis system for disease prediction using hybrid extreme learning machine. *International Journal of Intelligent Engineering & Systems*, 16(1): 1-12. <https://doi.org/10.22266/ijies2023.0228.01>
- [20] Rath, M., Jain, N., Bist, P., Agrawal, T. (2020). Smart healthcare model: An end-to-end framework for disease prediction and recommendation of drugs and hospitals. In *High Performance Vision Intelligence: Recent Advances*, pp. 245-264. Singapore: Springer Singapore. https://doi.org/10.1007/978-981-15-6844-2_17
- [21] Yu, Z., Huang, F., Zhao, X., Xiao, W., Zhang, W. (2021). Predicting drug-disease associations through layer attention graph convolutional network. *Briefings in Bioinformatics*, 22(4): bbaa243. <https://doi.org/10.1093/bib/bbaa243>
- [22] Rajesh, T.R., Rajendran, S. (2022). Intelligent multi-agent reinforcement learning based disease prediction and treatment recommendation model. In 2022 International Conference on Augmented Intelligence and Sustainable Systems (ICAIS), Trichy, India, pp. 216-221. <https://doi.org/10.1109/ICAIS55157.2022.10010747>
- [23] Nagaraj, P., Deepalakshmi, P. (2022). An intelligent fuzzy inference rule-based expert recommendation system for predictive diabetes diagnosis. *International Journal of Imaging Systems and Technology*, 32(4): 1373-1396. <https://doi.org/10.1002/ima.22710>
- [24] Lei, X., Fang, Z., Guo, L. (2019). Predicting circRNA-disease associations based on improved collaboration filtering recommendation system with multiple data. *Frontiers in Genetics*, 10: 897. <https://doi.org/10.3389/fgene.2019.00897>
- [25] Sharma, V., Samant, S.S. (2023). A multi-level decision-making framework for heart-related disease prediction and recommendation. *System Research and Information Technologies*, 4: 7-20. <https://doi.org/10.20535/SRIT.2308-8893.2023.4.01>
- [26] Yu, Z., Amin, S.U., Alhussein, M., Lv, Z. (2021). Research on disease prediction based on improved DeepFM and IoMT. *IEEE Access*, 9: 39043-39054. <https://doi.org/10.1109/ACCESS.2021.3062687>