GR JOURNALS | Journal of Information and Communications Technology: Algorithms, Systems and Applications

Volume 1 Issue 1 (June 2025)
DOI: https://doi.org/10.64189/ict.25305
J. Inf. Commun. Technol. Algorithms, Syst. Appl., 2025, 1(1), 25305

Research Article

OPEN ACCESS

Intelligent Formulation Recommendation System: Leveraging Ayurvedic Classical Texts for Disease-Specific and Pharmacologically Tailored Drug Suggestions

Kaustubh Rathod,* Devesh Rathi, Sankalp Naranje and Jayashri Bagade

Department of Information Technology Engineering, BRACT's Vishwakarma Institute of Information Technology, Pune, 411048, Maharashtra, India

*Corresponding Author
Kaustubh Rathod
kaustubh.22110323@viit.ac.in

Received: 05 April 2025
Revised: 30 May 2025
Accepted: 11 June 2025
Published Online: 13 June 2025.

Citation

K. Rathod, D. Rathi, S. Naranje, J. Bagade, Intelligent formulation recommendation system: leveraging ayurvedic classical texts for disease-specific and pharmacologically tailored drug suggestions, *Journal of Information and Communications Technology: Algorithms, Systems and Applications,* 2025, 1(1), 25305, https://doi.org/10.64189/ict.25305.

Abstract

The vast knowledge of Ayurveda on individual plants and formulas based on unique qualities is priceless, yet it is frequently impractical to access this treasure of knowledge. In order to make the process of choosing the best Ayurvedic formulations based on symptoms, patient characteristics, and contraindications easier, this research presents a custom software solution. With the goal of giving Ayurvedic practitioners and students a user-friendly platform, the software offers vital insights into the various facets of traditional medicinal texts, such as sources, synonyms, and pharmacological qualities. This intelligent programme aims to assist the Ayurvedic community in making well-informed and efficient healthcare decisions by tackling navigational and scatteredness concerns.

***Keywords*:** Medicine; Machine learning; Ayurveda; Random-forest-algorithm; Decision tree; Formulation recommendation system; Disease-specific suggestions; Healthcare decision making.

1. Introduction

An enormous amount of information has been gathered about the medicinal qualities of individual plants and their harmonious combinations in formulations by the ancient Indian healing system known as Ayurveda. However, because it can be difficult to sort through a huge number of dispersed and big-scale sources of knowledge, this tremendous reservoir of expertise remains mostly untapped. This study proposes a revolutionary intelligent formulation recommendation system that uses the power of old Ayurvedic literature to deliver pharmacologically customized and disease-specific medicine recommendations to address this problem. By providing a thorough and intuitive platform, the suggested software solution seeks to enable Ayurvedic practitioners and students to make well-informed decisions when choosing suitable Ayurvedic formulas. Through the integration of data from multiple sources, including contemporary research, clinical practice, and classical books, the system offers a comprehensive summary of the pharmacological characteristics, indications, contraindications, and therapeutic effects of different formulations stated by Kyalkond *et al.*[1] With the use of sophisticated recommendation algorithms and an extensive knowledge base, the system can offer appropriate formulations depending on the unique traits, symptoms, and contraindications of each patient. Additionally, the system offers a user-friendly interface that makes it easier to find pertinent information, addressing navigational issues that are common with traditional medicinal texts. Information regarding particular formulations, such as sources, synonyms, and pharmacological characteristics, can be easily accessed by users. This improved accessibility encourages evidence-based practice and a deeper comprehension of the fundamental ideas of Ayurvedic treatment stated in Paulson and Ravishankar.[2] To put it simply, the intelligent formulation recommendation system acts as a link between the practical requirements of students and practitioners of Ayurveda and the extensive knowledge found in classical Ayurvedic books of Risina Rasmith *et al*.[3] in Machine Learning-Based Detection System for Facial Skin Diseases and Ayurvedic Remedies. The system facilitates the Ayurvedic community's ability to make informed and effective healthcare decisions by simplifying the process of choosing suitable formulations and offering extensive insights into their therapeutic properties. This, in turn, advances Ayurvedic medicine.

The extensive body of knowledge on specific plants and formulas found in the Ayurvedic literature offers priceless insights into customary medical procedures. But getting to this wealth of information can be difficult at times, making it difficult for students and practitioners to fully utilize it. As a result, an increasing amount of study has looked into the creation of personalized software programs meant to make it easier to choose the best Ayurvedic formulations depending on a patient's symptoms, personal traits, and contraindications stated by Kale *et al*.[4] In order to shed light on the development and significance of intelligent programmes intended to improve the effectiveness of Ayurvedic healthcare decisions, this literature review aims to present an overview of the body of research that has already been done in this area. Basavaraj *et al* discussed different statistical features are retrieved for each signal and categorized using the K-NN classifier to identify three different types of Doha's.[5] Monitoring System is not portable /wearable and comparatively more expensive health monitoring system. The accuracy of the KNN Model is higher compared to other selected models for an experiment. CNN architecture based on AlexNet for classification of medicinal plants Image preprocessing to convert scanned images to 256x256x3 dimensions. Existing technologies were unable to emulate the different types of therapeutic plant species present in India. The CNN method can be made better by hyperparameter tweaking, data redesigning, and model optimisation stated by Hegde *et al*.[6]

Model-01 with BoVW and SVM outperforms all other datasets when compared with 94% accuracy on the newly constructed one. KNN is preferable over the support vector machine for this kind of application with 100% accuracy by Thella *et al*.[7] Using MATLAB tool R2019a, the accuracies for KNN were obtained at 100% and for SVM at about 93.23% by Roopashree *et al*.[8] The highest accuracy was gained by CNN Model. The KNN model gained the highest accuracy of 91.06%. Certain ML models have a parameter-dependent nature, hindering disease prediction accuracy. Some models have relatively low accuracy percentages for disease prediction stated by Raghukumar *et al*.[9] SVM achieves high accuracy levels for different categories of medicinal plants, ranging from 92.5% to 99.5%. AUC values higher than 0.9 suggest outstanding discrimination. Poor quality control, inappropriate herb substitutions, confusion in identification, and challenges in manual recognition of dried plants undermine the efficacy of Ayurvedic medicine, posing risks of incorrect usage and unpredictable side effects, highlighting the crucial need for strong quality control in the industry by Kalpana Joshi.[10] Dileep and Pournami studied Ayur-Vriksha and achieved a commendable classification accuracy of 97% based on a trained dataset containing more than 50 leaf samples of medicinal plants.[11] The model's utilization of Sanskrit words for plant identification adds an additional layer of cultural relevance. Despite the high accuracy, there are limitations to Ayur-Vriksha. The system's performance might be affected by variations in lighting conditions, and the accuracy may decrease when applied to a broader range of medicinal plant species not covered

in the training dataset.

The machine learning-based system successfully identifies four facial skin conditions (acne, dark circles, dark spots, and wrinkles) and recognizes 20 different Ayurvedic plants with high accuracy. The system's accurate detection of skin conditions, Ayurvedic plant recognition, and personalized remedies contribute to overall skincare. While there are challenges, the approach enhances patient engagement through a user-friendly web application and telemedicine system, paving the way for effective, technology-driven skincare solutions studied by Sharoni. Marques *et al.* predicted Ayurveda-based constituent balancing using machine learning faces challenges.[12] Limited and diverse datasets, the intricate nature of Ayurvedic principles, subjective diagnoses, external factors' influence, dynamic practices, ethical concerns, and integration with traditional methods pose potential limitations. These factors need careful consideration for the effective and responsible implementation of machine learning in Ayurveda were studied by Batvia *et al.*[13] Vinayak *et al* summarized model based on the Seq2Seq LSTM model with an attention mechanism achieved an optimum accuracy of 98.6% in generating summaries of Ayurvedic plant information.[14] The research concludes that the developed mobile-based application is capable of providing reliable and accurate information about Ayurvedic plants. The marker-based watershed algorithm and VGG-16 model were found to be the most suitable for object detection and classification, respectively.

2. Methodology

When creating an Intelligent Formulation Recommendation System using classical Ayurvedic texts, a methodical approach comprising multiple crucial stages is required. In order to give a fundamental understanding and identify gaps in current knowledge, a thorough assessment of the literature on Ayurvedic principles, classical texts (such as Charaka Samhita and Sushruta Samhita), and previous works connected to Ayurvedic recommendation systems is first conducted. The phases of the research process that follow are informed by the literature review phase. After the evaluation of the literature, gathering and compiling data becomes crucial. Reputable sources, traditional texts, and scholarly articles provide accurate information about ayurvedic medicines, formulations, qualities, therapeutic uses, contraindications, and interactions. In order to guarantee the validity and correctness of the data gathered, domain experts are essential as stated in Satish Nadiga *et al* Identification of Ayurveda Herbs using Machine Learning.[15] This stage entails carefully organizing and structuring the data to make knowledge extraction and computational analysis easier (Fig. 1).

The creation of a solid knowledge base that incorporates the gathered information in an organized manner follows. Relationships between various items in the Ayurvedic domain are mapped out using ontology-based modelling, which guarantees semantic consistency and interoperability. The foundation for later algorithm development and suggestion creation is provided by this knowledge base. A key component of the process is algorithm development, which entails building algorithms that can produce suggestions for tailored formulations based on input characteristics such as patient symptoms, constitution (Prakriti), disease diagnosis, and contraindications stated by Marada Srinivasa Rao *et al* in A Methodology for identification of Ayurvedic Plant based on Machine Learning Algorithm.[16] Ayurvedic formulations and their therapeutic efficacy for particular health disorders are correlated with patterns

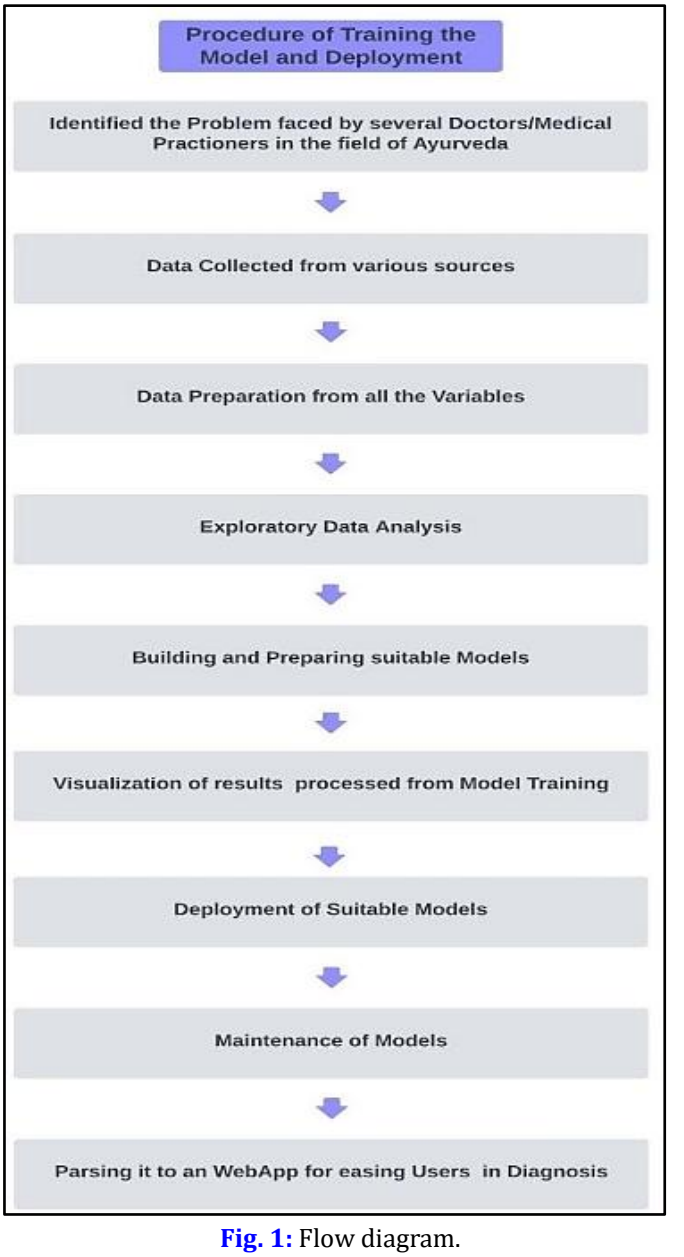

Fig. 1: Flow diagram.

and correlations found in machine learning approaches such as collaborative filtering and supervised learning. In order to guarantee adherence to Ayurvedic principles and guidelines during recommendation creation, rule-based reasoning techniques are also implemented. A proper dataset, including the disease names and the diagnosis for them, is compiled and trained. Dataset creation is the most tedious task in formulating proper results, as it needs to be validated by different health experts to make sure the results must imbibe correct medicine for the asked disease diagnosis.

The next stage entails developing an intuitive software interface that can be used on mobile or web platforms. This would allow students and Ayurvedic practitioners to enter patient data and get customized formulation recommendations instantly. The program includes features for perusing Ayurvedic texts, seeing comprehensive details on therapeutic herbs and formulas, and investigating associated ideas. To determine the developed system's accuracy, relevance, and usefulness, validation and assessment are essential. Ayurvedic practitioners and students participate in validation studies to provide input and evaluate the system's effectiveness. Iterative enhancements to the program are guided by metrics including user happiness, coverage of Ayurvedic texts, recommendation accuracy, and efficiency in supporting decision-making. These metrics are assessed. Throughout the research process, ethical issues such as transparency, data security, and privacy are carefully taken into account. Policies and procedures governing software development for the healthcare industry are followed, and precautions are taken to protect user and patient data. The dissemination and documentation of study findings are essential for adding to the body of knowledge in academia and encouraging more studies in this area. The approach, methods, software architecture, and validation outcomes are covered in depth in a research paper or technical report that is ready for presentation at pertinent conferences and seminars as well as publication in peer-reviewed publications. This guarantees the broad distribution of information and encourages cooperation and input from the scientific community, propelling ongoing development and progress in the area of Ayurvedic formulation recommendation systems stated by Pradeep Tiwari *et al*. in Recapitulation of Ayurveda constitution types by machine learning of phenotypic traits.[17]

2.1 Modeling and analysis

2.1.1 Data collection and preprocessing

Our recommendation engine is based on a vast collection of classical Ayurvedic books, including scholarly works, treatises, and old manuscripts. These books provide a goldmine of information regarding remedies, qualities, and the impact of medicinal plants on a range of illnesses. First, these documents had to be digitized and organized into a format that could be used for computer analysis. Tokenization, stemming, and lemmatization are a few text preprocessing techniques that were used to standardize and eliminate noise from the text. To improve the data's interpretability and usefulness, additional attempts were undertaken to connect terminologies to the corresponding botanical names and pharmacological characteristics.

2.1.2 Feature engineering

Our recommendation system's efficacy mostly depends on how well Ayurvedic ideas and formulations are represented. To convert unstructured textual data into understandable numerical representations, feature engineering was used. This required the use of methods like word embeddings, in which semantic links between words are captured by mapping them to high-dimensional vectors. To further capture the spirit of Ayurvedic principles, domain-specific elements like rasas (tastes), gunas (qualities), and doshas (biological energies) were retrieved and included in the feature space.

Min-Max Scaling: This technique reduces standard deviations and suppresses the impact of outliers on the feature by scaling the feature to a specified range, often between 0 and 1. where Xmax and Xmin are the maximum and minimum values of the feature, and x is the instance's individual value (person 1, feature 2).

Feature scaling

An additional method of normalization is to divide the feature by its range, which is represented as X_{max} - X_{min}, after deducting the minimal value, Xmin, from the feature. This provides us with:

$$XScaled = \frac{Xi - Xmin}{Xmax - Xmin} \quad (1)$$

The provided feature is mapped onto the interval [0,1] by this normalization.

Normalization

With the exception of replacing the minimum value with the mean value of the full set of data for each entry, this method is substantially the same as the previous one. The results are then divided by the difference between the minimum and maximum values.

$$XScaled = \frac{Xi - Xmean}{Xmax - Xmin} \quad (2)$$

Standardization

The primary basis of this scaling technique is the data's variance and central tendency. First, the data that needs to be normalised should have its mean and standard deviation ascertained. The next step is to subtract the mean value from each item and divide the result by the standard deviation. Assuming the data are already

normal but skewed, this helps achieve a normal distribution of the data with a mean of zero and a standard deviation of one.

$$XScaled = \frac{Xi - Xmean}{\sigma} \quad (3)$$

- Scaling

We employ two primary statistical metrics of the data in this scaling procedure. We are to divide the result by the interquartile range and subtract the median from each item after computing these two values.

$$XScaled = \frac{Xi - Xmean}{IQR} \quad (4)$$

2.1.3 Model selection and training

Several machine learning algorithms were explored to develop the recommendation system, each tailored to address specific aspects of the problem stated by Vani Rajasekar, Sathya Krishnamoorthi, Muzafer Saracevi ˇ c´, Dzenis Pepic, Mahir Zajmovic and Haris Zogic in Ensemble Machine Learning Methods To Predict The Balancing Of Ayurvedic Constituents In The Human Body. These algorithms include, but are not limited to:

- Collaborative Filtering: Leveraging user-item interactions and similarities between formulations to make personalized recommendations.
- Content-Based Filtering: Analyzing the intrinsic properties of formulations and matching them with user preferences and requirements.
- Hybrid Models: Combining collaborative and content-based approaches to leverage the strengths of both methodologies.

A combination of supervised and unsupervised learning methods were used to train the models. In supervised learning, predictive models were trained using past patient symptoms, features, and treatment results data. Unsupervised learning methods, including clustering, were used to find patterns and put related formulations in groups according to their characteristics and outcomes.

2.1.4 Evaluation metrics

We assessed the recommendation system's performance using common measures including F1-score, recall, accuracy, and precision. Metrics like mean average accuracy (MAP) and normalized discounted cumulative gain (NDCG) were also taken into consideration for personalised recommendation jobs to evaluate the ranking and relevancy of suggested formulations.

2.2 Mathematical formulations of machine learning algorithms

In the research, the Intelligent Formulation Recommendation System was developed using the Random Forest and Decision Tree algorithms. These algorithms were developed on historical data that included patient profiles, symptoms, and treatment outcomes by utilising classical Ayurvedic books. Through the integration of domain-specific characteristics like doshas, gunas, and rasas, the algorithms produced tailored and situation-specific suggestions for Ayurvedic formulas. Random Forest's ensemble approach guaranteed generalizability and robustness, whereas Decision Trees offered comprehensible insights into the decision-making process.

2.2.1 Random Forest Algorithm

During training, random forests (RF) build a large number of distinct decision trees. The final prediction, which is the mean prediction for regression or the mode of the classes for classification, is derived from the sum of the predictions made by all the trees. They are called ensemble approaches because they use a set of findings to arrive at a final judgment. The feature relevance is determined by multiplying the likelihood of accessing a node by the weighted decrease in impurity at that node. The number of samples that reach the node divided by the total number of samples yields the node probability. The more significant the trait, the higher its worth.

Gini Impurity:

A metric used to assess a dataset's impurity, especially in decision tree nodes, is the Gini impurity. It determines the probability of a wrong classification based on the dataset's class distribution, assuming a randomly selected sample is labelled. The following is the formula for Gini impurity: "Gini impurity *(Gini(p))* is calculated by subtracting the sum of squared probabilities of each class (p_i) from 1, where *i* ranges over all classes in the dataset."

$$Gini(p) = 1 - \sum_{i=1}^{J} P_i^2 \quad (5)$$

Here, p_i represents the probability of an element belonging to class *i* and J, represents the total number of classes.

Information Gain: In decision tree methods, information gain is a metric used to evaluate how well a dataset is split depending on a specific attribute. It quantifies the split's reduction in entropy or chaos. The following is the formula for information gain: "Information gain *(IG(D, f))* is obtained by subtracting the entropy of the dataset (H(D)) from the conditional entropy of the dataset given a feature *(H(D|f))*."

$$IG(D, f) = H(D) - H(D|f) \quad (6)$$

In this case, *H(D)* denotes the entropy of dataset *D, IG(D, f)* is the information gain of dataset *D* with regard to feature *f*, and *H(D|f)* denotes the conditional entropy of dataset *D* given feature *f*.

Bootstrap Sampling

Bootstrap sampling is a technique used in Random Forest to create multiple datasets for training decision trees. It involves randomly selecting samples from the original

dataset with replacements. Each sample is of the same size as the original dataset. The probability of selecting a particular data point in each sampling is *1/N*, where *N* is the size of the original dataset.

- Out-of-Bag (OOB) Error: Out-of-Bag error is an estimation of the model's performance using samples not included in the bootstrap samples for each tree. The expected proportion of out-of-bag samples for each tree is approximately 1/e, where e is Euler's number. The OOB error is calculated by evaluating the model's performance on these out-of-bag samples.
- Voting (Classification): In classification tasks, the Random Forest combines the predictions of multiple decision trees by majority voting. Each tree predicts a class for a given sample, and the final predicted class is the one with the most votes among all trees.

Random Forest is a key component in the Intelligent Formulation Recommendation System, which analyzes patient data and classical Ayurvedic texts to suggest appropriate formulations based on symptoms, patient features, and contraindications. The system makes use of numerous decision trees, each of which was trained using a subset of the attributes that were taken from the patient profiles and textual data. Through the consolidation of these trees' predictions, the system may offer context-aware, individualized medication recommendations that cater to the specific requirements of each patient.

2.2.2 Decision tree algorithm

A well-liked and adaptable supervised learning technique for both regression and classification applications is the decision tree algorithm. Since it is non-parametric, it does not assume anything about the distribution of the underlying data. Decision trees are constructed by recursively dividing the feature space into regions (or leaves) according to the values of input features. This allows the trees to be optimised for information gain or impurity reduction at each split.

- Entropy: An indicator of dataset impurity is entropy. The entropy *H(S)* of a set *S* with proportion *p* of samples labelled as class *1* and *1–p* classified as class *0*, for a binary classification problem with classes 0 and 1, is given by:

$$H(S) = - p \log_2(p) - (1-p) \log_2(1-p) \quad (7)$$

- Information Gain: When a dataset is divided based on a specific feature, the amount of entropy (impurity) that is reduced is measured. The Information Gain is computed as follows given a dataset *D* with *N* samples and *K* classes and a feature *A* with potential values *{a1,a2,..., am}*, the information gain is calculated as:

$$IG(DA) = H(D) - \sum_{i=1}^{m} |D_i|/D|H(D_i)/ \quad (8)$$

Where,
H(D) is the entropy of dataset *D*, $|D_i|$ is the number of samples in *D* for which feature *A* has value a_i, and *H(Di)* is the entropy of the subset of *D* for which feature *A* has the value *ai*.

- CART (Classification and Regression Trees) Cost Function: The CART algorithm usually minimizes impurity (either entropy or Gini impurity) when splitting decision trees used in classification and regression applications. The mean squared error, or MSE, is frequently utilized as the cost function in regression.
- Classification Prediction: A decision tree classification model moves through the tree from the root node to a leaf node in order to forecast a given sample based on the feature values of the sample. The expected class is the majority class of the training data in the leaf node.
- Prediction in Regression: In decision tree, regression tasks are predicted by regression models, which take the average of the goal values of the training samples in the leaf node—a node that can be reached by ascending the tree from the root node. This is predicated on the feature values of the sample.

The Intelligent Formulation Recommendation System uses decision trees to analyze and extrapolate meaning from patient data and classical Ayurvedic texts. The system can discover pertinent elements and their interactions by building decision trees. This allows the system to offer suitable formulations based on the given symptoms, patient characteristics, and contraindications. Decision trees give practitioners and students insight into the decision-making process and make it easier for them to comprehend the reasoning behind each proposal.

3. Results and discussion

Technology is here to stay, and it is up to us to make the most of it. Ayurvedic principles and practices can be seamlessly integrated with the newest technologies, thanks to several developments that have emerged in recent years.

3.1 Dataset description

The description of all the diseases and symptoms and the ayurvedic medicines for the respective ones have been taken from the textbook, which was according to the syllabus of the Central Council of Indian Medicine, New Delhi. The dataset purely reflects all the symptoms that the human body faces in day-to-day life. A dataset is divided into attributes, namely age, sex, symptoms, diseases, and ayurvedic medicines for respective diseases. References have been added for all the ayurvedic prescriptions to make a practitioner aware of where the results have been fetched. The study started with compiling and digitizing traditional Ayurvedic manuscripts and carefully extracting insightful information on therapeutic herbs, formulas, and their pharmacological characteristics. To provide customized

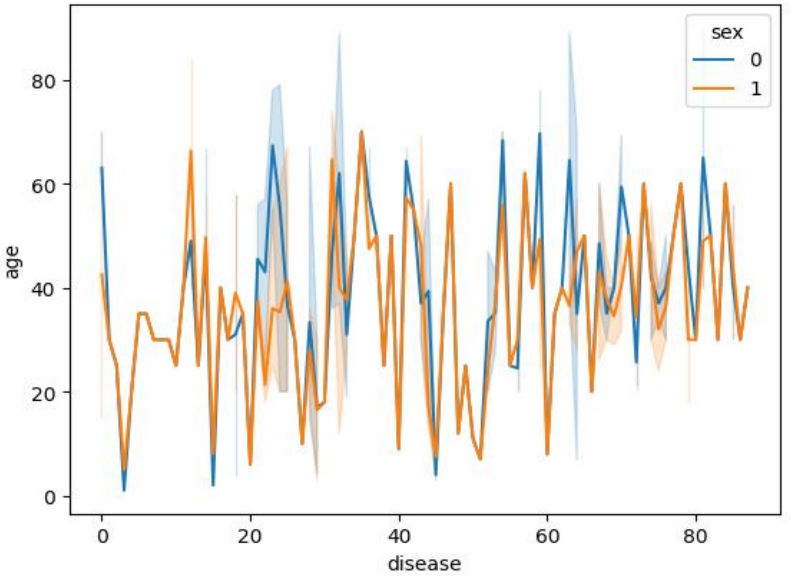

Fig. 2: Age vs Disease plot.

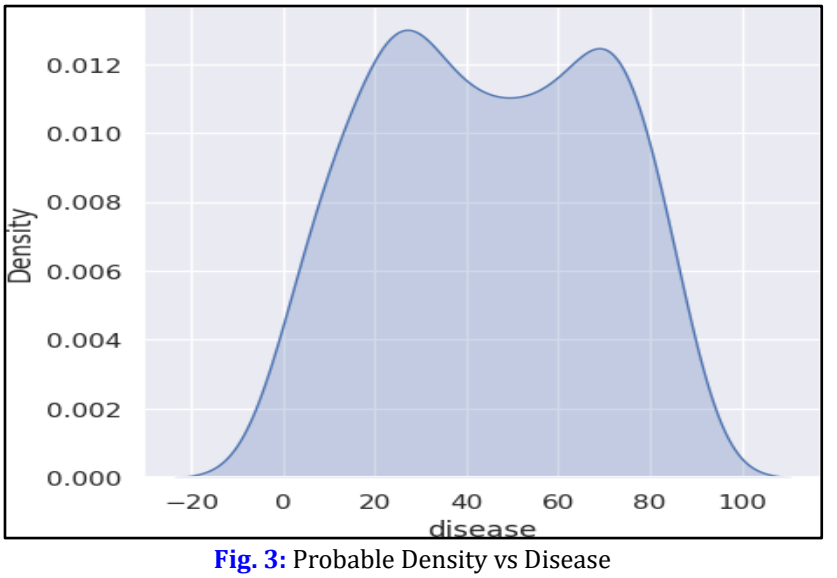

Fig. 3: Probable Density vs Disease

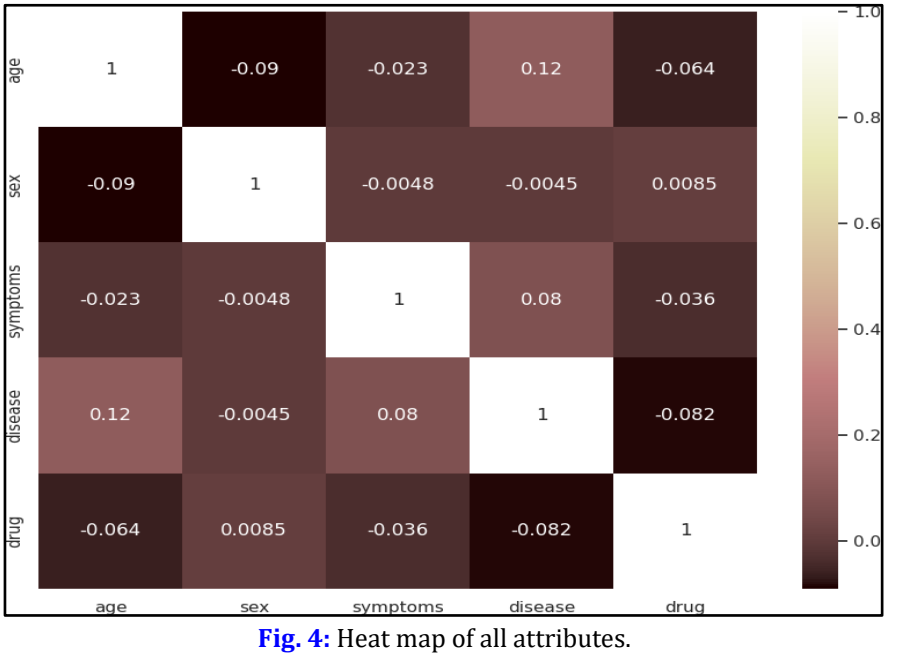

Fig. 4: Heat map of all attributes.

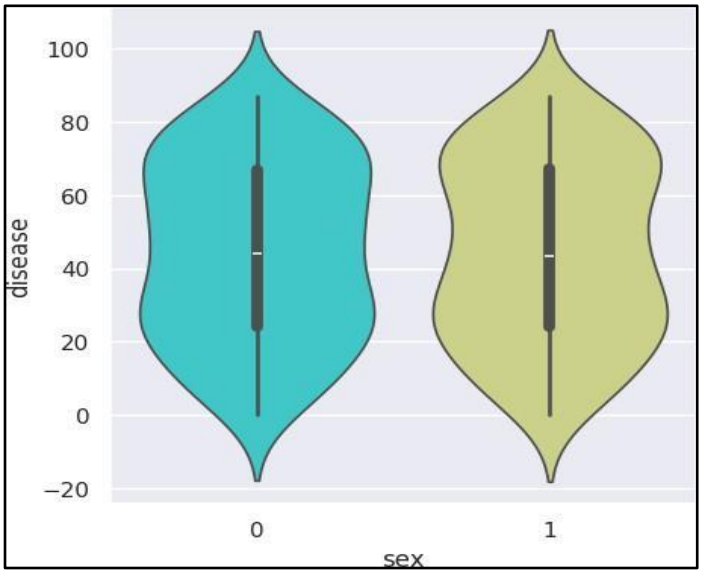

Fig. 5: Trend graph of Diseases vs Sex.

medication recommendations, patient information on symptoms, traits, and contraindications was also gathered.

3.2 Graphs and Plots

3.2.1 Age vs Disease Graph

The age distribution of the patients was analyzed, and the results provided fascinating information on the prevalence of the disease in various age groups. Fig. 2 highlights the need to take age into account when formulating recommendations by showing a higher frequency of particular diseases in particular age groups.

3.2.2 Probable Density vs Disease plot

Gender-specific patterns in disease occurrence were shown by a comparative examination of diseases selected for men and women. One can visually assess any relationship between the probability density being studied and the occurrence of the disease. The plot in Fig. 3 highlights the need for specialized healthcare interventions by illuminating differences in disease prevalence between genders.

3.2.3 Heat map of all attributes

A heat map Fig. 4 was created to illustrate the relationship between several attributes, such as symptoms, traits, and contraindications. This thorough depiction makes it easier to see how different features relate to one another, which promotes a more comprehensive comprehension of patient profiles.

3.3 Trends of diseases in male and females

Notable variations in disease prevalence and patterns were found by analyzing the diseases that were selected specifically for men and women. Fig. 5 shows different health profiles for men and women were indicated by the gender-specific patterns that several diseases showed in our sample. For example, disorders like heart disease and problems connected to the prostate were more common in men, which may be attributed to physiological variations and lifestyle choices unique to gender. On the other hand, women showed greater rates of autoimmune diseases and reproductive health issues, highlighting the impact of hormonal variations and genetic predispositions. Comprehending the distinct disease patterns associated with gender is crucial in customizing healthcare treatments and treatment approaches to accommodate the disparate requirements of males and females. Healthcare professionals can maximize patient care and improve health outcomes for both genders by acknowledging and addressing these inequalities.

3.4 Calculations and tabular differences

A key component in the creation of the intelligent formulation recommendation system is the mathematical computations that underpin the machine learning model training process. These computations are essential for turning raw data from patient profiles and old Ayurvedic books into useful insights for tailored medication recommendations. The first step is featuring engineering, in which unstructured data is organized into a feature matrix X, where each row corresponds to a patient or formulation and each column represents a particular feature, such as patient attributes, symptoms, warning signs, and pharmacological properties of formulations.
The prediction given by this model serves as a greater convenience for all doctors and practitioners to cure diseases and give medications according to the symptoms input. The model takes Age, Sex, and symptoms as input and gives medicine and their references as output shown in Table 1.

Meaningful data representations are filled into the feature matrix by use of the transformation function $f_j(x_i)$.

Table 1: Prediction table.

Example	Age	Sex	Symptoms	Disease	Predicted Output
1	30	M	Sneezing	Common Cold	Devadaru
2	84	M	Fatigue	Arthritis	Boswellia, Ginger, and Turmeric

Table 2: Parametric differences.

Model	Precision	Accuracy	F1 - Score
Random Forest Model	92.08	93.33	92.61
Decision Tree Algorithm	70	70	70

The best Ayurvedic medications are then predicted using machine learning models trained on this feature matrix X, such as random forest and Decision Trees. Using optimization algorithms like gradient descent, the model's parameters θ are iteratively updated in order to reduce the difference between the actual and predicted drug suggestions, which is measured by a loss function $J(\theta)$. Through this iterative process, the system is able to improve its prediction powers and, in the end, provide individualized, evidence-based medication recommendations that cater to the unique characteristics consideration. Two different types of Algorithms/Models predicted the medicines, viz., the Random Forest Model and the Decision Tree Algorithm. The dataset is extensively trained and tested using both the models and inferences have been found in Table 2. The intelligent formulation recommendation system, which uses machine learning algorithms to provide practitioners with individualized and scientifically supported medication recommendations, represents a paradigm shift in Ayurvedic treatment. Through the integration of age, gender, and attribute correlations into formulation recommendations, the method improves patient care quality and moves Ayurvedic medicine closer to evidence-based practice and better patient outcomes.

4. Conclusion

The results obtained from the intelligent formulation recommendation system signify a significant advancement in leveraging Ayurvedic classical texts for personalized healthcare decision-making. Through the integration of machine learning algorithms, the system offers tailored recommendations based on patient-specific parameters, thereby enhancing the efficacy and efficiency of Ayurvedic treatment approaches. The analysis of age distribution among patients revealed age-specific trends in disease prevalence, underscoring the importance of age consideration in formulation recommendations. Additionally, insights derived from the density of disease plots provide valuable guidance for prioritizing healthcare interventions based on disease burden. Furthermore, the gender-specific analysis highlights the need for gender-tailored healthcare approaches, recognizing distinct disease patterns among men and women. This understanding enables practitioners to deliver personalized care that accounts for gender-specific nuances in disease occurrence. Using extensive data analysis and machine learning methods, this research initiative tackles the long-standing problem of gaining access to and utilizing the abundance of Ayurvedic knowledge in clinical practice. The system's easy-to-use interface and tailored recommendations enable students and Ayurvedic practitioners to make informed and effective healthcare decisions, improving treatment outcomes and patient care. The dissemination and documentation of study findings are essential for adding to the body of knowledge in academia and encouraging more studies in this area. The approach, methods, software architecture, and validation outcomes are covered in depth in this research. This guarantees the broad distribution of information and encourages cooperation and input from the scientific community, propelling ongoing development and progress in the area of Ayurvedic formulation recommendation systems.

Funding Declaration

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

Data Availability Statement

The datasets generated and/or analyzed during the current study that support the findings are available from the corresponding author upon reasonable request.

Conflict of Interest

There is no conflict of interest.

Artificial Intelligence (AI) Use Disclosure

The authors declare that artificial intelligence (AI)-assisted tools were used only for language refinement, grammar improvement, and manuscript structuring purposes during the preparation of this work. All technical content, experimental implementation, results, and interpretations were independently developed and verified by the authors.

Supporting Information

Not applicable.

References

[1] S. A. Kyalkond, S. S Aithal, V. M. Sanjay, P. S. Kumar, A novel approach to classification of ayurvedic medicinal plants using neural networks, *International Journal of Engineering Research & Technology*, 2022, **11**, doi: 10.17577/IJERTV11IS010128.

[2] A. Paulson, S. Ravishankar, AI based indigenous medicinal plant identification, 2020 Advanced Computing and Communication Technologies for High-Performance Applications (ACCTHPA), Cochin, India, 2020, 57-63, doi: 10.1109/ACCTHPA49271.2020.9213224.

[3] R. K. A. Risina Rasmith, K. G. Chamindu Hansana, C. P. Abeywickrama, S. Siriwardana, H. L. D. P. De Silva, S. Jayaweera, Machine learning-based detection system for facial skin diseases and ayurvedic remedies, *International Journal of Innovative Science and Research Technology*, 2023, 8, doi: 10.5281/zenodo.10066306.

[4] S. G. Kale, S. Jain, S. Rangari, R. C. Dharmik, M. Gardi, V. S. Lande, Identification of medicinal leaves and recommendation of home remedies using machine learning, *International Journal of Intelligent Systems and Applications in Engineering*, 2024, **12**, 407–413.

[5] K. Basavaraj, S Balaji, Nadi Pariksha: A novel machine learning based wrist pulse analysis through pulsauscultation system using K-NN classifier, *International Journal of Electronics and Communication Engineering and Technology*, 2021, **12**, 1-10, doi: 10.34218/IJECET.12.3.2021.001.

[6] P. L. Hegde, A. Harini, A Text Book of Dravyaguna Vijnana, According to the syllabus of Central Council of Indian Medicine, New Delhi, 2nd ed, Chaukhamba Publication.

[7] P. K. Thella, V. Ulagamuthalvi, A comparative analysis on machine learning models for accurate identification of medical plants, *Revista Geintec*, 2021, **11,** doi: 10.47059/REVISTAGEINTEC.V11I4.2309.

[8] S. Roopashree, J. Anitha, Enrich Ayurveda knowledge using machine learning techniques, *Indian Journal of Traditional Knowledge*, 2020, **19**, 813-820, doi: 10.56042/ijtk.v19i4.44537.

[9] A. M. Raghukumar, G. Narayanan, Comparison of machine learning algorithms for detection of medicinal plants, 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2020, 56-60, doi: 10.1109/ICCMC48092.2020.ICCMC-00010.

[10] K. Joshi, Leveraging artificial intelligence as a tool to improve health services and modernize ayurveda treatment-a perspective, *Journal of Research in Ayurvedic Sciences*, 2023, **7**, S10-S12, doi: 10.4103/jras.jras_85_23.

[11] M. R. Dileep, P. N. Pournami, AyurLeaf: A Deep Learning Approach for Classification of Medicinal Plants," TENCON 2019-2019 IEEE Region 10 Conference (TENCON), Kochi, India, 2019, 321-325, doi: 10.1109/TENCON.2019.8929394.

[12] O. Marques, Integrating contemporary technologies with Ayurveda: Examples, challenges, and opportunities, 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Kochi, India, 2015, doi: 10.1109/ICACCI.2015.7275809.

[13] V. Batvia, D. Patel, A. R. Vasant, A survey on ayurvedic medicine classification using tensor flow, *International Journal of Computer Trends and Technology*, 2017, **53**, 68-70, doi: 10.14445/22312803/IJCTT-V53P114.

[14] V. Majhi, B. Choudhury, G. Saha, S. Paul, Development of a machine learning-based Parkinson's disease prediction system through Ayurvedic dosha Analysis, *International Journal of Ayurvedic Medicine*, 2023, **14**, 180–189, doi: 10.47552/ijam.v14i1.3228.

[15] S. Nadiga, Bindu, Jyotishri, Veenaxi Painginkar and Vinoliya Sharline Pinto, Identification of ayurveda herbs using machine learning, *International Research Journal of Modernization in Engineering Technology and Science*, 2023, **05**, 2071-2074, doi: 10.56726/IRJMETS34656.

[16] M. S. Rao, S. P. Kumar, K. S. Rao, A methodology for identification of ayurvedic plant based on machine learning algorithm, *International Journal of Computing and Digital Systems*, 2023, 14, doi: http://dx.doi.org/10.12785/ijcds/140196.

[17] P. Tiwari, R. Kutum, T. Sethi, A. Shrivastava, B. Girase, S. Aggrawal, R. Patil, D. Agrawal, P. Gautam, A. Agrawal, D. Dash, S. Ghosh, S. Juvekar, M. Mukerji, B. Prasher, Recapitulation of ayurveda constitution types by machine learning of phenotypic traits, *PLOS One*, 2017, **12**, e0185380.doi: 10.1371/journal.pone.0185380.