AI-Powered Ayurvedic Doshas Assessment Tool: Health and Nutrition for all.

ABHINANDAN P. MANGAONKAR[1], STEVEN NORONHA[2], SADIYA RAJGOLI [3], OMKAR KALLLEKAR[4], PRAJWAL ARKASALI[5]

1 Department of Computer Science Engineering, SG Balekundri Institute of Technology, Belagavi 590001, India
Email: abhimangaonkar123@gmail.com
2 Department of Computer Science Engineering, SG Balekundri Institute of Technology, Belagavi 590001, India
Email: stevennoronha9@gmail.com
3 Department of Computer Science Engineering, SG Balekundri Institute of Technology, Belagavi 590001, India
Email: sadiyarajgoli@gmail.com
4 Department of Computer Science Engineering, SG Balekundri Institute of Technology, Belagavi 590001, India
Email: omkark6342@gmail.com
5 Department of Computer Science Engineering, SG Balekundri Institute of Technology, Belagavi 590001, India
Email: Prajwalarkasali007@gmail.com

ABSTRACT

This paper introduces a groundbreaking Ayurvedic dosha prediction tool developed using Artificial Intelligence Markup Language (AIML) and advanced machine learning techniques such as decision trees, gradient boosting, and random forest methods. Through meticulously designed structured questionnaires, this innovative tool evaluates individual characteristics and behaviors, providing precise determination of the predominant dosha constitution. It further offers personalized diet plans aligned with Ayurvedic principles tailored to the identified dosha type, fostering optimal health and balance. Additionally, users can effortlessly locate nearby Ayurvedic doctors and clinics through an integration feature, facilitating access to professional healthcare services. By seamlessly blending AIML technology with sophisticated machine learning methods, this tool transcends conventional boundaries, offering a comprehensive solution for Ayurvedic health management. It empowers users to seamlessly integrate holistic wellness practices into their daily lives, promoting harmony between mind, body, and spirit. This represents a significant advancement in Ayurvedic health technology, bridging modern science with ancient wisdom to enhance overall well-being and promote a balanced lifestyle.

Keywords:

Ayurveda, Dosha prediction, Machine learning, Holistic wellness, Artificial intelligence.

1. INTRODUCTION

His holiness Ashtavaidyan Thrissur Thaikat Mooss, a famous figure in South Indian Ayurvedic tradition, explained Prakruti Dosha in his book "Ayurveda Elohim" as follows: "Prakruti is the true nature of the doshas, the persons born with; which influences not only the physical but also mental, emotional characteristics as well".

[1] Vata Dosha

[2] Pitta Dosha

[3] Kapha Dosha

On the other hand, we have the so-called "Vikruti", meaning the current imbalance of doshas from the unique

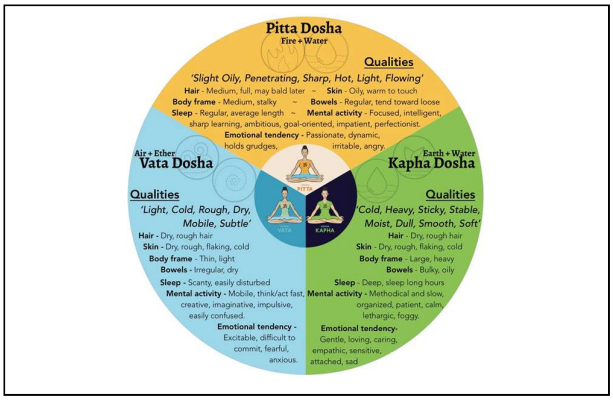


Fig 1: Types of ayurvedic Prakruti

balance of doshas. However, in Ayurveda, not only the body but everything in the universe is made up of Panchamahabhutas, the five elements - Space, Air, Fire,

Water, and Earth in different compositions. Different bodies have their own Prakruti Dosha depending on the said unique compositions and Prakruti represents the beginning of life. This is never being changed and remains the same throughout an individual's life from birth to death. Maintaining the given area and understanding this unique balance is the very key to prevent diseases and promote healthy wellbeing.

Trying to find out the Prakruti Dosha of an individual is indeed significant because every human being is believed to be born with a unique balance of doshas. Depending on this unique balance, the body works accordingly. This is why Prakruti Dosha is sometimes referred to as "what a man is born with" and in reality, Prakruti represents the genome. In modern terminology, Prakruti Dosha can be considered as the unique genetic code but very simply.

Doshas are the energies which make up an individual's physical and mental characteristics. Vata, Pitta, and Kapha doshas are three of these which signify the five elements - Air, Fire, Water, Earth, and Space. How an individual's body works and responds to both internal and external changes is highly affected by these three doshas. These are individually

known as - Vata Dosha from Air and Space, Pitta Dosha from Fire and Water, and Kapha Dosha from Earth and Water.

Prakruti Dosha is a fundamental and unique characteristic of Ayurveda and this represents the primordial and self-born quality of doshas. The significance of identifying Prakruti dosha in Ayurvedic medicine is emphasized by the fact that this is necessary for a proper understanding of the balance of an individual's body, mind, and spirit at a given time. Cultivating and maintaining this balance through an in-depth understanding of Prakruti Dosha forms the basis of Ayurvedic treatment.

Ayurveda is considered one of the world's oldest holistic healing systems. According to this ancient medical approach, all the things that exist in the universe - both living and nonliving - are connected. Besides, health and wellness depend on a delicate balance between the mind, body, and spirit. The primary goal of Ayurveda is to maintain this balance and prevent the occurrence of illness; and to do so, it focuses on what is known as Prakruti Dosha.

1.1 Ayurveda and Prakruti Dosha

In Ayurveda, the five fundamental elements of the universe are space, air, fire, water, and earth, which compose the three principles or 'doshas' in the body. These doshas are Vata, Pitta, and Kapha. Prakruti refers to the basic constitution of a person, depicting the genetic and prenatal factors. This is the combination of the three doshas, present at the time of conception and stays constant throughout life. It is considered as the physical and mental state of the body. An individual's Prakruti gives a detailed picture of the basic physical and psychological characteristics and limitations of that person. This helps the physician in deciding the clinical diagnosis, planning for treatment, and other options of healthcare in the most effective manner. Also, it is very helpful in understanding the prognosis of the different types of diseases and also in recognizing the healthy status of a patient. It is a highly valuable concept in Ayurveda, not found in Western medicine. On the other hand, the term 'Vikruti' refers to the current state of the doshas in the body. Vikruti is the change in the doshas from the time of your birth to the present state. A person's lifestyle, diet, work, exercise, and all the other activities can influence the vikruti of the body. In simple words, if Prakruti is the constitutional diagnosis, then Vikruti is the pathological diagnosis. Vikruti of a person can be varying at different stages of life and with different factors like seasonal changes, stress, diet, and lifestyle changes, which can affect the doshas. Ayurveda considers the patient as a whole, which includes body, mind, and soul. Because of the unique concepts of the Prakruti and Vikruti, Ayurveda has established a special method of clinical diagnosis, called 'Darsanam Parva,' which is nothing but systematic observation. Based on the observation, questioning, and some of the modern methods like pulse diagnosis, tongue diagnosis, eye diagnosis, and the body constitution examination, the Ayurvedic physician can confirm the Prakruti, Vikruti, and the nature of the diseases. Thus, Prakruti is a fundamental basis of the physical and mental characteristics of a person and important in clinical diagnosis and Ayurvedic treatment. On the other hand, Vikruti explains the current state of the doshas in the body. In clinical practices, the physician uses the 'Prakruti Pariksha,' which means the examination of the body constitution and 'Nadi Pariksha,' which is the pulse diagnosis, to confirm the Prakruti of a person. According to Ayurveda, the three doshas have the capacity to influence an individual's mental and physical states. When in balance, Vata results in creativity and vitality, Pitta in intelligence, and Kapha in love and forgiveness. However, when not in balance, Vata can cause fear and anxiety, Pitta can cause frustration and anger, and Kapha can cause attachment and greed. It is said that for a perfect and happy life, a person should always be active and take up many mental and physical activities. There should be a right balance between the three doshas, and a person's lifestyle and diet also reflect that right balance. Also, it shows how to treat chronic diseases and how the diseases can be avoided. Doshas are constantly aggravated and deranged by different factors, and close contact with the seasons appears to be the most significant one. Also, the derangement of 'Agni,' which means the digestive fire, can be considered as the root of all pathologies. Ayurveda highlights the importance of balancing Prakruti, Vikruti, and Agni in the quest for optimal health.

1.2 Definition and Concept of Prakruti Dosha

Prakruti, a Sanskrit word, means "nature" and Dosha denotes "mind body type." Prakruti is the unique constitution of each individual and is determined at the time of conception. From the Ayurvedic point of view, disease is a state of imbalance, disharmony, or discomfort within the individual as a whole and Prakruti is an essential aspect in Ayurvedic diagnosis. The Prakruti body-type is determined primarily by three elements of our constitution known as the "Doshas". Vata Dosha is a combination of space and air, pitta of fire and water, and kapha of water and earth. Every individual has Vata, Pitta, and Kapha Doshas in their body, but the predominance of each Dosha determines the individual's Prakruti. Vata Prakruti individuals are etheric, thin, quick, alert, creative, affectionate, and lively. Pitta Prakruti individuals tend to be sharp, medium build, compact, muscular, faster mind but with less retention. As for Kapha Prakruti individuals, they are slower in movement and slower in speech, but with a good memory once they managed to learn. This kind of Dosha is more related to physical body characteristics and hereditary and it is more stable than Dosha imbalance. However, one or more Dosha can be emphasized depending on the lifestyle, diet, or any emotional or physical trauma. Prakruti shows our basic nature, our natural inborn equilibrium. Vata, Pitta, and Kapha are the true Doshas and are called the Prakruti Doshas. Doshas form the basis of Ayurvedic medicine, and the three Doshas respectively reflect the three basic forms of life energy: movement, heat transformation, and lubrication and structure. They support the body's functions while the elemental activities occur and the physical materials gain nourishment from the digestive processing, and each Dosha has a preferred site for providing functionality. The Doshas here not only indicate the physiological functions of the body but also useful in diagnosis and treatment and maintenance of health as well.

By understanding Doshas, it is much easier to balance one's life and enhance health by choosing favorable foods and lifestyle to counteract the unfavorable dietary and habitual excess. Doshas reflect the body's intelligence. Dosha accepts mental and even spiritual physicality of the individual. The Prakruti Dosha - The Nature of what one is. Every treatment is given according to the need and condition of the patient. It makes Ayurvedic medicine applicable to all ages and conditions. Dosha is very extensive; it guides the Ayurvedic physician in diagnosis but also prognostic sign and treatment alternatives. It is the most vivid, powerful diagnostic therapeutic tool in the application of Ayurvedic medicine. Doshas can either show the doctor what the disease is and also show in which way the disease will shape. Doshas represent what lies within the genetic code of the individual and reflect the individual state of the Doshas/Prakruti and the state of Disturbed Dosha/Vikruti. By understanding Prakruti and Vikruti, one's condition and preventive measures can easily be established. Though the condition of Prakruti and Vikruti will change with time, what is important is to return the patient to what they were born and live with this nature. Last but not least, every individual is influenced by their own Prakruti nature, but the penetration of the senses and mind also plays a critical role in the physical and mental welfare. Even with the same life experience, the feeling and reflection will be different - past experience reacts to present experience and the behavior alters our future. However, Ayurveda conveys the nature that enables one to live to a fine age and still retain freedom from disease. Doshas provide a very comprehensive approach to maintain health, to restore and renew balance, and even to prevent acute illness emerging. By understanding people's Doshas, people will learn to support and promote self-healing.

1.2.1 Types of Prakruti Dosha

According to Ayurvedic philosophy, the body is governed by three different doshas. These are called Vata, Pitta, and Kapha. Each dosha is a combination of two of the five elements. Vata is a combination of space and air, Pitta is fire and water, and Kapha is earth and water. The doshas are very important in Ayurvedic science and form the basis of all diagnosis and treatment. They are used to both describe body types and identify illnesses.

Every individual has all three doshas in their body, just in different amounts. However, usually one dosha will be stronger than the others, which is an individual's prakruti - their body and mental constitution. People with predominantly Vata doshas are generally quite thin, with a slender frame. They are quick thinking and are usually quite fast at doing things. However, they can easily become stressed and have restless minds. When out of balance, Vata sufferers can often suffer from flatulence, constipation, and joint pain. On the skin, they will develop rashes and dry patches.

Pitta types have a medium, muscular body and are strong and active. They are very precise in their thoughts and like to be in control. However, they can sometimes become irritable and possibly angry under stressful conditions. Pitta imbalances can lead to inflammations, rashes, ulcers, and acne.

Kapha Prakruti people are usually blessed with good health, have a strong immune system, and are naturally quite happy. When out of balance, they can become overweight, suffer from fluid retention, and sinus problems. Kapha type individuals will enjoy routine and find it difficult to adjust to change. They have a much slower pace than Vata types.

It is important to remember that an individual's Prakruti is the balance of the three doshas at the moment of conception. It does not change during one's lifetime. However, the doshas making up the Prakruti may become imbalanced due to poor lifestyle, stress, poor diet, and toxins. This is when illnesses develop. Every person is different and needs to keep their body and mind in balance, which reflects the underlying principle of Ayurvedic medicine.

1.3 Assessment and Diagnosis of Prakruti Dosha

In recent years, there has been growing interest in integrating machine learning techniques, such as decision tree classifiers, into the realm of Ayurvedic diagnosis. Decision tree classifiers are algorithms that recursively partition the feature space into subsets based on the values of input features, ultimately leading to a prediction or classification.In the context of Ayurveda, decision tree classifiers could be trained on datasets containing information about patients' Prakruti doshas along with various physiological, psychological, and lifestyle factors. These algorithms could then learn to identify patterns and relationships among these variables to predict an individual's Prakruti dosha.The advantage of using decision tree classifiers lies in their ability to handle complex datasets and capture non-linear relationships between variables. By leveraging large datasets, these algorithms can potentially uncover subtle correlations between genetic markers, environmental factors, and Prakruti doshas that may not be immediately apparent to human practitioners.However, the effectiveness of decision tree classifiers in Ayurvedic diagnosis ultimately depends on the quality and representativeness of the data used for training. Additionally, the interpretability of the resulting models is crucial in order for practitioners to trust and understand the predictions made by these algorithms.Integrating decision tree classifiers into Ayurvedic practice represents a fusion of ancient wisdom with modern technology, offering the potential to enhance diagnostic accuracy and provide personalized healthcare recommendations based on an individual's unique constitution and lifestyle. However, further research and validation are needed to fully realize the potential of these computational approaches in the context of Ayurveda.

1.4 Methods and Tools for Prakruti Assessment

Finally, it takes a look at the application of personalized medicine— Prakruti-based medicine— in the modern era. The section on assessment and diagnosis methods of Prakruti Dosha is particularly lengthy and in-depth. The following instruments and methodologies are employed, "in one form or another," worldwide to assess Prakruti: 1) Self-assessment questionnaires 2) Physician's clinical assessment 3) Prakruti analytical software 4) Phenotype markers-based assessment 5) Sibling-pair analysis from genome-wide association

research. The paper provides a comprehensive insight into each of the methods and their aim and limitation. It discusses the evolution of Prakruti assessment methodologies from mere clinical observation to the use of recent technological tools, thereby accentuating the paradigm shift in Prakruti assessment. Not only is the theory behind each of the methods described in detail but also the flow of the actual assessment. The practicality of the methods in different situations, together with the merits and drawbacks of the methodologies, have also been touched on. This paper, once again, evidences the fact that Ayurveda is not just a traditional knowledge but a constantly evolving science with the adoption of new tools and technologies. It also shows the seriousness of the research in Prakruti and its potential implication in personalized medical development. Overall, the main idea for this section is to showcase the advance of technology in Prakruti assessment. The section tries to bring the readers from the past, on how ancient physicians do observation and documentation, to the future, on how phenotype markers and genome-wide association research will pave the way for modern Prakruti-based medicine. In the meticulous assessment and diagnosis of Prakruti Dosha, a multifaceted and systematic approach is essential, integrating advanced methodologies such as decision tree algorithms for comprehensive analysis. This methodological framework encompasses a series of intricate steps, beginning with the administration of meticulously crafted questionnaires to individuals seeking Prakruti assessment.

These questionnaires are designed to capture a broad spectrum of data points, spanning various dimensions that influence an individual's Prakruti. From physiological attributes such as body constitution and metabolic tendencies to lifestyle habits, dietary preferences, environmental exposures, and even psychosocial factors, the questionnaires aim to gather a holistic portrait of the individual's constitution and inherent tendencies.Once the questionnaires have been completed and the data collected, the process advances to the analytical phase, where the power of decision tree algorithms comes into play. Decision tree algorithms, renowned for their ability to discern complex patterns and relationships within datasets, are employed to meticulously analyze the vast array of information gathered from the questionnaires.

Through the intricate traversal of decision tree branches, practitioners can uncover subtle correlations, identify dominant traits, and ascertain nuanced nuances within an individual's Prakruti profile. This analytical process not only aids in understanding the unique composition of an individual's constitution but also serves as a guiding framework for personalized treatment strategies.By harnessing the insights gleaned from the application of decision tree algorithms, practitioners can tailor treatment approaches that resonate with the individual's inherent nature and health needs. From dietary recommendations and lifestyle modifications to herbal formulations and therapeutic interventions, the personalized treatment strategies derived from Prakruti assessment are designed to address the individual's specific imbalances and promote holistic well-being.

Table 1: Characteristics of tri-doshas

Body, hair Color	Eye Color	Eye Size	Hair Growth
Hair Type	Nails color	Palms Color	Scalp hair Color
Skin Color	Skin Nature	Teeth Color	Appetite Amount
Bowel Freq	Bowel Tendency	Digestive Amount	Sleep Amount
Sleep Quality	Stool Consistency	Walking Amount	Anger Quality
Making friends speed	Retaining friend quality	Hair feel	Skin cracked
Skin freckle	Skin mark	Skin pimple	Skin wrinkled

The psychosomatic constitution of an individual can be categorized into seven factors known as VATA, PITTA, KAPHA, VATA-PITTA, VATA-KAPHA, PITTA-KAPHA, and balanced (SAMDOSHA). It's common for individuals to have dominance in more than one factor, placing them within specific dosha types. Ayurvedic science classifies individuals based on their physical and physiological characteristics. Understanding one's prakriti aids doctors in guiding their patients toward a healthier lifestyle. Prakriti is influenced by various factors including genetics, prenatal and postnatal influences. Analyzing prakriti assists in understanding the body's needs, promoting overall health across personal, familial, and professional domains.

In essence, the integration of decision tree algorithms into the assessment of Prakruti Dosha represents a paradigm shift towards precision and personalized medicine within the realm of Ayurveda. It embodies a commitment to understanding each individual's unique constitution and harnessing this knowledge to optimize health outcomes and foster a harmonious balance between body, mind, and spirit

2 Methodology

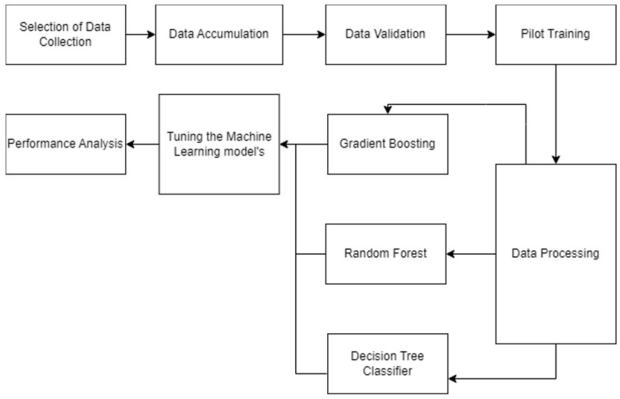


Fig 2: Flow Diagram

A. Data Collection:

The initial step involves comprehensive data collection from individuals, focusing on a wide array of factors including health status, lifestyle choices, dietary habits, emotional tendencies, and physical attributes. This data is typically gathered through detailed questionnaires designed to capture nuanced information relevant to Ayurvedic principles and the concept of Prakriti. Each individual's responses serve as crucial features for the machine learning model, forming the foundation for personalized health assessments. Additionally, our approach includes a thorough examination of case papers from Ayurveda colleges, where

we collect and analyze detailed patient records. By incorporating real-world data from clinical settings, we enrich our dataset with insights from Ayurvedic practitioners, ensuring a comprehensive understanding of individual health profiles and enhancing the accuracy of our predictive models.

B. Data Accumulation:

Accumulating a diverse and representative dataset is crucial for building an effective Prakriti detection model. This involves gathering data from various sources, including individuals of different ages, genders, ethnicities, and geographical locations, to ensure the model's applicability across diverse populations. Additionally, incorporating data from individuals with varying health conditions and lifestyles enables the model to capture a comprehensive range of factors influencing Prakriti. Collaborations with healthcare institutions, wellness centers, and research organizations can facilitate access to relevant datasets, while ensuring compliance with ethical and privacy regulations. By accumulating high-quality and diverse data, the model can be trained to accurately predict Prakriti types and provide personalized health insights to individuals worldwide.

C. Data Validation:

Ensuring the quality and consistency of the collected data is paramount for building an accurate predictive model. This entails meticulous data cleaning procedures to eliminate duplicates, handle missing values, and address outliers that could potentially skew the analysis. Additionally, categorical variables are transformed into numerical representations suitable for machine learning algorithms. Exploratory Data Analysis (EDA) techniques are employed to gain insights into the distribution of features and identify any anomalies or patterns within the dataset. Moreover, we collaborated with Dr. Usharani, an esteemed Ayurvedic practitioner, who provided validation and guidance throughout the data validation process. Dr. Usharani's expertise and insights enhanced the rigor of our validation procedures, ensuring that our dataset accurately reflects the principles of Ayurveda and the concept of Prakriti.

D. Pilot Training:

Before deploying the Prakriti detection model at scale, conducting a pilot training phase is essential to assess its performance and feasibility in real-world scenarios. During this phase, the model is trained on a subset of the data to evaluate its effectiveness in accurately predicting Prakriti types. The pilot training also allows for fine-tuning of model parameters and validation of the data collection process to ensure its suitability for the intended application. Feedback from pilot users is collected and incorporated into model refinement efforts, enabling iterative improvements before full-scale deployment. By conducting thorough pilot training, potential issues can be identified and addressed early, ultimately enhancing the model's accuracy and usability when deployed for broader use.

E. Data Preprocessing:

Data preprocessing plays a pivotal role in preparing the raw dataset for effective utilization in building the Prakriti detection model. This phase involves a series of steps aimed at cleaning, transforming, and enhancing the quality of the data to optimize model performance. Initially, data cleaning techniques such as handling missing values, removing duplicates, and addressing outliers are applied to ensure data integrity. Subsequently, categorical variables are encoded into numerical representations suitable for machine learning algorithms. Feature scaling or normalization may also be performed to bring features to a similar scale, preventing any single feature from dominating the model training process. Additionally, feature selection or dimensionality reduction methods may be employed to reduce the computational complexity and enhance the model's efficiency. By meticulously preprocessing the data, the model can learn more effectively from the input features, resulting in improved accuracy and robustness in predicting Prakriti types.

F. Gradient Boosting:

Gradient Boosting is an ensemble learning technique that sequentially combines multiple weak learners, typically decision trees, to create a strong predictive model. In the context of Prakriti detection, Gradient Boosting can effectively capture complex relationships between the input features and the individual's Prakriti type. By iteratively minimizing the errors of the previous models, Gradient Boosting produces a robust ensemble model that can accurately classify individuals into their respective Prakriti categories.

G. Random Forest:

Random Forest is another ensemble learning method that constructs multiple decision trees during training and outputs the mode of the classes (classification) or mean prediction (regression) of the individual trees. In the context of Prakriti detection, Random Forest offers several advantages, including handling high-dimensional data with ease, mitigating overfitting, and providing insights into feature importance. By aggregating the predictions of multiple decision trees, Random Forest can deliver accurate and reliable predictions of an individual's Prakriti constitution.

H. Decision Tree Classifier:

Decision tree is a non-parametric classifier and a predictive model based on the divide and conquer strategy It is a classic example of soft computing and solves the purpose of the classifier. A decision tree consists of the top node as root, lower child nodes, branches and internal nodes in the context of Prakriti detection, Decision Trees provide interpretable models that can directly represent the decision-making process. Each branch of the tree corresponds to a decision based on a particular feature, leading to a clear understanding of how the model arrives at its predictions. While Decision Trees may lack the complexity of ensemble methods like Gradient Boosting and Random Forest, they offer transparency and ease of interpretation, making them valuable for understanding the underlying factors contributing to an individual's Prakriti type.

I. Fine-tuning the machine learning model:

It is a critical step in optimizing its performance for accurate Prakriti detection. Model tuning involves adjusting hyperparameters and optimizing various aspects of the

algorithm to achieve the best possible predictive accuracy. For Gradient Boosting, tuning parameters such as learning rate, tree depth, and the number of estimators can significantly impact model performance. By carefully selecting these hyperparameters through techniques like grid search or random search, we can find the optimal combination that minimizes overfitting and maximizes predictive power. Similarly, Random Forest models require tuning of parameters such as the number of trees, maximum depth of trees, and the minimum number of samples required to split a node. Adjusting these parameters helps strike a balance between model complexity and generalization ability, ensuring robust predictions across different datasets.

Decision Trees also benefit from parameter tuning, with key parameters including the maximum depth of the tree, minimum number of samples required to split a node, and the minimum number of samples required at each leaf node. By tuning these parameters, we can prevent overfitting and create decision trees that generalize well to unseen data.

Moreover, techniques like cross-validation are employed to evaluate the model's performance across different parameter settings and ensure its robustness. By systematically adjusting hyperparameters and validating the model's performance, we can fine-tune the machine learning model to achieve optimal accuracy and reliability in predicting an individual's Prakriti type. Overall, model tuning is a crucial aspect of building an effective Prakriti detection model, enabling us to harness the full potential of machine learning algorithms and deliver personalized health assessments based on Ayurvedic principles.

J. Performance Analysis of Machine Learning Models:

Evaluating the performance of the machine learning models used in the Prakriti detection system is vital to ensure accurate and reliable predictions of an individual's Prakriti type. Each model, including Gradient Boosting, Random Forest, and Decision Trees, undergoes thorough performance analysis to assess its effectiveness in predicting Prakriti constitutions.

For Gradient Boosting, performance analysis involves evaluating metrics such as accuracy, precision, recall, and F1-score. By fine-tuning hyperparameters like learning rate and tree depth, we aim to achieve the optimal balance between model complexity and generalization ability. Additionally, techniques like ROC curves and precision-recall curves provide insights into the model's ability to discriminate between different Prakriti types and its trade-offs between precision and recall.

Similarly, Random Forest models undergo performance analysis to assess their predictive accuracy and robustness. By tuning parameters like the number of trees and tree depth, we aim to optimize model performance while mitigating overfitting. Performance metrics such as accuracy and area under the ROC curve (AUC-ROC) provide valuable insights into the model's discriminative power and ability to generalize to unseen data.

For Decision Trees, performance analysis focuses on evaluating metrics such as accuracy, precision, and recall. By tuning parameters like tree depth and minimum sample split, we aim to create decision trees that capture the underlying patterns in the data while avoiding overfitting. Visualization techniques such as confusion matrices and decision tree plots aid in interpreting the model's predictions and understanding its decision-making process.

Overall, performance analysis of the machine learning models used in the Prakriti detection system ensures their effectiveness in accurately predicting Prakriti constitutions. By systematically evaluating performance metrics and fine-tuning model parameters, we can deliver personalized health assessments based on Ayurvedic principles with high accuracy and reliability.

2.1 Meal plan and Diet Recommendations for Each Dosha

Our system offers personalized meal plans tailored to each user's specific Dosha type, aligning with the principles of Ayurveda as outlined by Ayurvedic doctors. Pitta types are recommended to favor cool, heavy, and dry foods, while avoiding hot, spicy, and oily dishes to maintain balance. Vata individuals are advised to opt for warm, cooked, and moist foods, steering clear of cold, bitter, and astringent foods to support their constitution. Similarly, Kapha types are encouraged to prefer warm, light, and spicy foods, while avoiding cold, heavy, and oily dishes to promote equilibrium within their body. The proposed diet recommendations are consistent with the characteristics associated with each Dosha, emphasizing the importance of diet and lifestyle in restoring the balance of Doshas and promoting overall health. By providing personalized meal plans based on Ayurvedic principles, our system aims to support individuals in achieving optimal well-being by addressing the underlying causes of imbalance and promoting holistic health. This personalized approach not only enhances physical health but also contributes to mental and spiritual well-being, offering individuals a comprehensive path to wellness rooted in ancient wisdom and tailored to modern lifestyles.

3. MACHINE LEARNING MODELS

A. Decision tree classifier

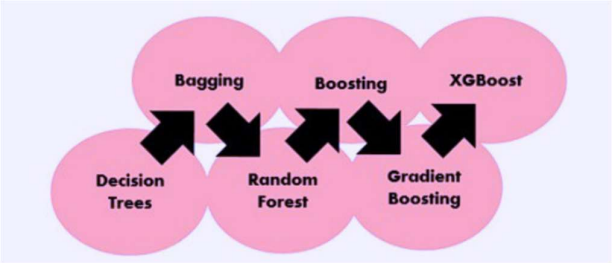
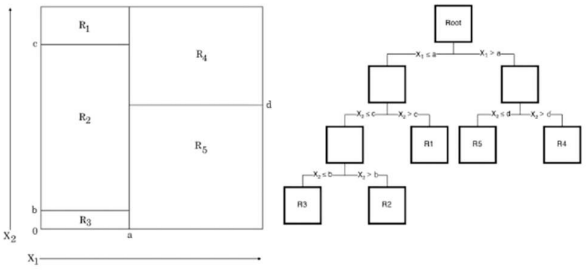


Fig 3: Evolution of decision tree

Decision Trees (DT): Decision Trees are fundamental models that recursively split data based on feature thresholds. A decision tree is a classifier. Decision trees are a widely used

technique in statistical learning, where they are constructed to fit an existing set of data, and then used to predict outcomes on new data. We start with N labeled "training records" of the form $(\boldsymbol{X},Y)$ where $\boldsymbol{X}$ is a k-dimensional vector of features describing the data we have, and Y is a label we give this record. For example, our data may be characteristics about humans and our label may be their shoe size.

Each component of $\boldsymbol{X}$ is called as "input variable", Y is called "dependent variable" or "target variable", and each row in such a table is called a "training example". Let us consider two input variables, such that $\boldsymbol{X}=(X1, X2)$. Further, let's assume there is an interesting value of $X1$ that we can split the dataset around and few interesting values of $X2$. Then, an example partitioning of our space of $(X1, X2)$ values is depicted in the left side of Figure 2, and a decision tree corresponding to such a partitioning is shown in the right side of the figure. Given an unlabelled vector $\boldsymbol{X} = (X1, X2)$, we first test whether $X1>a$. Then, if that turns out to be true, we test whether $X2>d$. This allows us to classify $\boldsymbol{X}$ into the region $R4$ or the region $R5$. If we initially had that $X1\leq a$, then we will test $X2$ against c and then against b, which allows us to further classify $\boldsymbol{X}$ into one of the regions $R1$,$R2$, or $R3$.



A B

fig. 4 a) Partition of 2D space in five regions b) Classifying decision tree into one of the five regions

Next, let Y take on a single constant value for each of the regions $R1,..,R5$. Let Yi be the value chosen for the region Ri, and let $Ii(\boldsymbol{X})$ be an indicator function that equals 1 when $\boldsymbol{X}\in Ri$. This allows us to obtain a model that can predict Y based on $\boldsymbol{X}$:

$$\hat{Y}(x) = \sum_{i=1}^{5} Y * li(X)$$

Obtaining such a model is the ultimate goal of training a decision tree. Same as the model represented in Fig. 2 as a partition of 2D space and as a decision tree.

The most basic process of training a decision tree on a dataset involves the following elements as,

- The selection of attribute
- Splits in the tree
- Stop splitting a node and mark it terminal
- The assignment of a label to each terminal node

i. Bagging (Bootstrap Aggregating): Bagging aims reduce variance by combining multiple Decision Trees. It involves randomly sampling subsets (with replacement) from the training data, training a Decision Tree on each subset, and aggregating predictions (e.g., majority vote for classification or average for regression). Random Forests are an ensemble of Decision Trees created using Bagging, improving robustness and generalization.

ii. Boosting: Boosting focuses on improving model accuracy by sequentially training weak learners. Weak learners are models slightly better than random guessing, and boosting combines their predictions to create a strong ensemble. Key concepts include Adaptive Boosting (AdaBoost), Gradient Boosting, and XGBoost (Extreme Gradient Boosting), which optimizes the gradient boosting process with regularization and parallelization.

iii. Gradient Boosting: Gradient Boosting builds an ensemble of weak learners (usually Decision Trees) by iteratively minimizing a loss function (e.g., mean squared error) and correcting errors made by previous models. XGBoost, an optimized version of gradient boosting, features regularization, parallelization, and tree pruning, making it widely used in Kaggle competitions and real-world applications.

iv. Application: Ensemble methods like Random Forests, Gradient Boosting, and XGBoost are powerful tools for classification, regression, anomaly detection, and ranking tasks such as recommender systems and search engines. In summary, ensemble methods combine weak models to create robust and accurate predictors, and understanding their evolution helps in choosing the right tool for specific tasks in machine learning.

B. Random forest algorithm

Random forest is a commonly-used machine learning algorithm, trademarked by Leo Breiman and Adele Cutler, that combines the output of multiple decision trees to reach a single result. Its ease of use and flexibility have fueled its adoption, as it handles both classification and regression problems.

Random forest algorithms have three main hyperparameters, which need to be set before training. These include node size, the number of trees, and the number of features sampled. From there, the random forest classifier can be used to solve for regression or classification problems.

The random forest algorithm is made up of a collection of decision trees, and each tree in the ensemble is comprised of a data sample drawn from a training set with replacement, called the bootstrap sample. Of that training sample, one-third of it is set aside as test data, known as the out-of-bag (oob) sample, which we'll come back to later. Another instance of randomness is then injected through feature bagging, adding more diversity to the dataset and reducing the correlation among decision trees. Depending on the type of problem, the determination of the prediction will vary. For a regression task, the individual decision trees will be averaged, and for a classification task, a majority vote—

i.e. the most frequent categorical variable—will yield the predicted class. Finally, the oob sample is then used for cross-validation, finalizing that prediction.

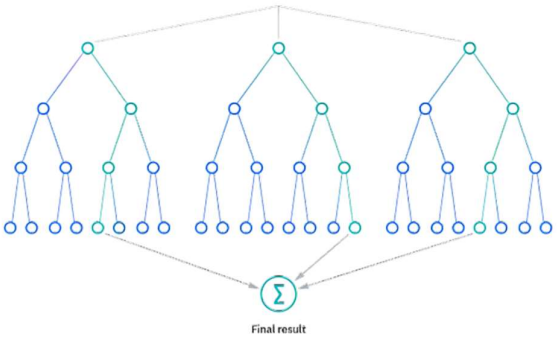


Fig 5: Random Forest Representation

C. Gradient Boosting algorithm

Gradient Boosting is an ensemble learning technique that combines multiple weak learners to form a strong learner. It is a powerful technique for both classification and regression tasks. Commonly used gradient boosting algorithms include XGBoost, LightGBM, and CatBoost. Each algorithm uses different techniques to optimize the model performance such as regularization, tree pruning, feature importance, and so on. Gradient Boosting is a prominent technique for boosting. In gradient boosting, each prediction corrects the inaccuracy of its previous. Unlike Adaboost, the weights of the training instances are not changed; instead, each predictor is trained using the predecessor's residual mistakes as labels.

Gradient Boosted Trees is a method whose basic learner is CART (Classification and Regression Trees).

The graphic below illustrates how gradient boosted trees are trained for regression situations.

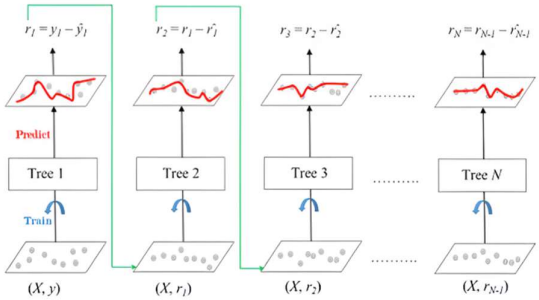


Fig 5: Gradient boosted trees representation

The ensemble is made up of N trees. The feature matrix X and the labels y are used to train Tree1. The y1(hat) predictions are utilized to calculate the training set residual errors r1. Tree2 is then trained with Tree1's feature matrix X and residual errors r1 as labels. The projected r1(hat) values are then utilized to calculate the residual r2. The technique is continued until all N trees in the ensemble have been trained.

This approach employs an essential parameter called as shrinkage.

Shrinkage refers to the fact that after multiplying the prediction of each tree in the ensemble by the learning rate (eta), which varies from 0 to 1, the forecast of each tree in the ensemble is shrunk. There is a trade-off between eta and the number of estimators; a decrease in learning rate must be compensated by an increase in estimators in order to achieve a specific level of model performance. Predictions may now be made because all trees have been taught. Each tree predicts a label, and the formula provides the final forecast.

y(pred) = y1 + (eta * r1) + (eta * r2) + ... + (eta * rN)

Gradient Boosting Regressor is the Scikit-Learn class for gradient boosting regression. Gradient Boosting Classifier is a classification algorithm that uses a similar approach.

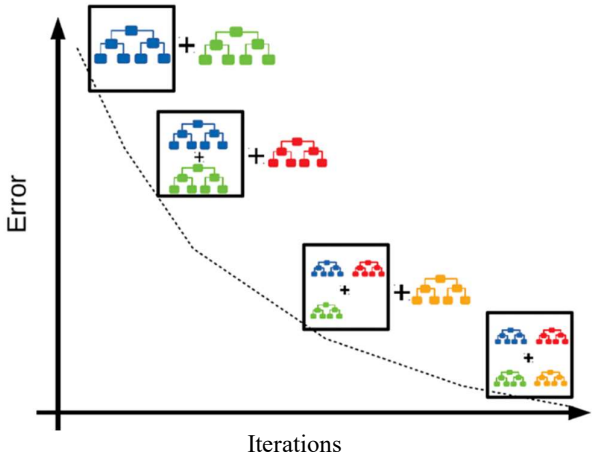


Fig 5: Gradient boost classification

4. RESULT

In this section, we describe the implementation of ML models on a pre-processed data. It is important to analyze the performance of different learning techniques consistently as in machine learning no single algorithm works best for every problem. There are many factors that makes an impact, such as the size and structure of your dataset. This is the reason behind the implementation of multiple algorithms for our data, while using a hold-out ‘‘test set’’ of data to evaluate performance and select the winner. Using traditional ML model, we are not finding any satisfactory results. For further improvement of the results, we apply hyper parameter tuning

A. SELECTION OF BEST LEARNING MODEL

It’s not possible to design optimal model architecture only by looking at the data set. Each model is based on different performance characteristics. We can estimate the accuracy of each model on unseen data values by using cross-validation. These estimates provide a path to choose best models among all referred models. It is a good idea to visualize the data using different techniques from different perspectives. Similarly, various ways are used for model selection.

The performance of different models was assessed using various metrics including precision, recall, F1-score, and support to gauge their accuracy. Figures 6, 7, and 8 display the evaluation results for the Decision tree, Random Forest,

and Gradient Boost models, respectively. In the context of our findings, it is important to note that in our classification results, the class labels correspond as follows: '0' signifies 'vata', '1' denotes 'kapha' and '2' represents 'pitta'.

	precision	recall	f1-score	support
0	0.69	0.82	0.75	11
1	0.85	0.92	0.88	12
2	1.00	0.79	0.88	14
accuracy			0.84	37
macro avg	0.85	0.84	0.84	37
weighted avg	0.86	0.84	0.84	37

Fig 6: Decision Tree classifier results

	precision	recall	f1-score	support
0	1.00	1.00	1.00	11
1	1.00	0.92	0.96	12
2	0.93	1.00	0.97	14
accuracy			0.97	37
macro avg	0.98	0.97	0.97	37
weighted avg	0.97	0.97	0.97	37

Fig 7: Random forest results

	precision	recall	f1-score	support
0	0.79	1.00	0.88	11
1	1.00	0.67	0.80	12
2	0.87	0.93	0.90	14
accuracy			0.86	37
macro avg	0.88	0.87	0.86	37
weighted avg	0.89	0.86	0.86	37

Fig 8: Gradient Boost results

CONCLUSION

In conclusion, the fusion of Artificial Intelligence with Ayurveda has yielded remarkable advancements. By harnessing sophisticated machine learning techniques, including decision trees, gradient boosting, and Random Forest, alongside cutting-edge AIML technology, innovative tools have emerged. These tools accurately predict dosha constitutions, offer personalized dietary recommendations, and enhance access to professional healthcare services. Among these techniques, Random Forest stood out, demonstrating superior performance and yielding the most promising results. This amalgamation of traditional Ayurvedic wisdom with modern technological capabilities not only empowers individuals to embrace personalized wellness practices but also bridges the gap between ancient traditions and contemporary healthcare. Through AI-driven solutions, individuals embark on a transformative journey towards optimal health and inner balance, guided by the timeless principles of Ayurveda. Thus, fostering a harmonious alignment of mind, body, and spirit amidst today's fast-paced world.

REFERENCES

[1] Predicting Ayurveda-Based Constituent Balancing in Human Body Using Machine Learning Methods," VISHU MADAAN AND ANJALI GOYAL., date of current version April 17, 2020., Digital Object Identifier 10.1109/ACCESS.2020.2985717

[2] Ayurveda: A Brief Introduction and Guide. Accessed:Jun. 23, 2019. [Online]. Available: https://www.ayurveda.com/resources/general-information

[3] E. Sadrfaridpour, T. Razzaghi, and I. Safro, ''Engineering fast multilevel support vector machines,'' Mach. Learn., vol. 108, no. 11, pp. 1879–1917,Nov. 2019

[4] A. Z. Woldaregay, E. Årsand, S. Walderhaug, D. Albers, L. Mamykina,T. Botsis, and G. Hartvigsen, ''Data-driven modeling and prediction of blood glucose dynamics: Machine learning applications in type 1 diabetes,'' Artif. Intell. Med., vol. 98, pp. 109–134, Jul. 20

[5] Possible Measures to Assess Functional States of Tridosha: A Critical Review Vandana Verma1, Sonam Agrawal2, Sangeeta Gehlot3 Department of KriyaSharir, Faculty of Ayurveda, Institute of Medical Science, Banaras Hindu University.

[6] R. Chinthala, S. Kamble, A. S. Baghel, N. N. L. Bhagavathi, ''Ancient

[7] archives of Deha-Prakriti (human body constitutional traits) in ayurvedic literature : A critical review,'' Int. J. Res. Ayurveda Pharmacy, vol. 10,no. 3, pp. 18–26, May/Jun. 2019

[8] Reliability of self-reported constitutional questionnaires in Ayurveda

[9] Diagnosis Corina Dunlap , Douglas Hanes , Charles Elder , Carolyn Nygaard , Heather Zwickey Accepted 24 April 2017 Original Research Article.

[10] Kurande VH, Waagepetersen R, Toft E, Prasad R. Reliability studies of diagnostic methods in Indian traditional Ayurveda medicine: an overview.J Ayurveda Integr Med 2013;4:67e76.

[11] Svoboda RE. Ayurveda, life health and longevity. New Delhi, India: Penguin Books; 1992.

[12] Khalsa KPS, Tierra M. The way of Ayurvedic herbs. Twin Lakes, WI: Lotus Press; 2008.

[13] 7. Thaker SJ, Gandhe PP, Godbole CJ, Bendkhale SR, Mali NB, Thatte UM, Gogtay NJ. A prospective study to assess the association between genotype, phenotype and Prakriti in individuals on phenytoin monotherapy. Journal of Ayurveda and integrative medicine. 2017 Jan 1;8(1):37-41.https://doi.org/10.1016/j.jaim.2016.12.001

[14] . V. Madaan and A. Goyal, "Predicting AyurvedaBased Constituent Balancing in Human Body Using Machine Learning Methods," in IEEE Access, vol. 8, pp. 65060-65070, 2020, doi:

10.1109/ACCESS.2020.2985717

[15] . Banerjee S, Biswas TK, Chattopadhyay K, Arzoo

[16] SH, Chattopadhyay B. An Approach to Screen Genotoxic-Susceptible Diabetic Population of Various Prakriti Groups for Personalized Disease Management. J Altern Complement Med. 2020 Oct 19. https://doi.org/10.1089/acm.2020.0001