**A Novel Approach on Ayurveda Dosha Prediction**

**On Big Data Environment**

**Mentorored by - Mr. Ashutosh Shankhdhar**

**Shashank Gupta** *(*[***shashank.gupta\_cs18@gla.ac.in***](mailto:shashank.gupta_cs18@gla.ac.in)*)*

**Achal Paliwal (**[***achal.paliwal\_cs18@gla.ac.in***](mailto:achal.paliwal_cs18@gla.ac.in))

**Department of Computer Engineering and Application , GLA University Mathura India.**

**Abstract**: On the basis of eternal substances Air, Water, Earth, Fire, Space Ayurveda has classified the human body in three physical humors or Doshas *VATA, KAPHA, PITTA* this principle is also known as Tri-Dosha. Every human body constituent is dominant for one “*Prakriti*” by birth while there are some people whose body is dominant for more than one prakriti. Studies of these body energies (Doshas) in Ayurveda has been used for a long time but the proper measurements and assurance of this principal (Tri-Doshas) still lags behind. Imbalance Doshas can be balanced by using proper diet. To get effective and meaningful dataset, 54 validated Characteristics of human body are used. Computations Algorithms used are Decision Tree, Support Vector Machine (SVM), K-Nearest Neighbor (KNN) and Naïve Bayes. For analyzing the performance of algorithm evaluation measures like root mean square (RMSE), Recall, Precision, F1-Score and accuracy are used. On analyzing the performance of algorithms on training and test data authors find that the data is best fitted using Decision Tree. Which gives 0.95 Precision, 0.78 Recall and 83.4 % Accuracy. After prediction of dosha the model will suggest the proper diet and some do and don’ts to balance the imbalanced doshas.

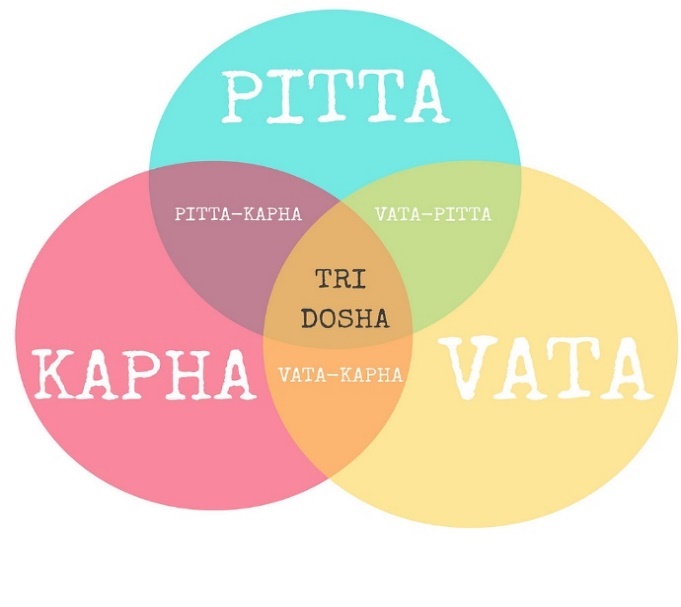
**I Introduction:** Ayurveda is considered as oldest healing Science (5000 Years) [1]. The word Ayurveda is combination of two words “Ayur” means *life* and veda means *know* that’s why also known as “*The Science of Life*” [2]. Ayurveda considered as the ancient and oldest tradition system of medicine [TSM’S] [2]. The equality between someone’s nature and self is the base upon which the principles of Ayurveda is built. It helps in maintaining healthy life by keeping person’s soul, mind, diet and body in balance. Just as every person has a unique fingerprint, each person has a particular pattern of energy an individual combination of mental, physical and emotional characteristics which comprises their own constitution. Ayurveda describes every body is dominant for one prakriti by birth there are some whose body predominant for more than one prakriti. Theses prakriti’s or elemental humours are known as *Tri-Dosha* [3] the three-principal constituent are VATA (movement and change), KAPHA (structure and fluidity) and PITTA (transformation and metabolism). In this digital and high-tech civilization every person is having diseases and imbalance diet due to this body leads to illness and in terms of Ayurveda it is described as *vikruti* “*The Nature of the Imbalance*”.

Figure 1. Tri-Dosha on bile nature as per Ayurveda.

**A**. **Tri-Doshas and Prapancha:** The five basic elements of life are *Space*, *Water, Fire*, *Air* and *Earth*, present in the human body every human body have physical and behavioural differences that are classified into one or more metabolic forces or *doshas* also know as tridoshas – vata, pitta & kapha [4]. Vata dosha is derived from the elements of Space(ether) and Air. Pitta dosha is derived from the elements of Fire and Water. And kapha dosha is derived from the elements of the Earth and Water [3].

**Table 1**. Tri-Doshas and Prapancha (Panchmahabhutas)

|  |  |  |
| --- | --- | --- |
| *Ekdoshaja* one dosha dominant | *dwjindoshaja* two-dosha dominant | Tridoshaja three doshas pre-dominant |
| VATA: Earth +Air | VATA+PITTA | VATA+PITTA+KAPHA |
| KAPAHA: Earth+Water | KAPHA+VATA | - |
| PITTA: Fire+Water | PITTA+KAPHA | - |

Human nature *(prakruti)* has basically two types: the physical nature *(Shaaririk Prakriti)* and the intellectual nature *(Manas Prakruti)* [1]. *VATT*, *PITT* and *KAPH* are *Sharirik prakriti. VATA,* prakruti associated with movement and change of body. Those with the vata dosha have body physique usually described as slim, energetic, and creative [5]. vata take care of all types of excretions, breathing and musculoskeletal functions. *PITTA*, prakriti associated with transformation and metabolism of body. PITTA prakriti person have keen digestion, moderate sleep, easily put on weight. KAPHA also pronouns as “*KUFFA”* prakrati associated with Structure and fluidity. KAPHA prakriti person has big bodies, dryness, moist hairs physique. These three humorous or dosha maintains all physical and mental processes in the living beings. The characteristics that describe the difference between these doshas are shown in table. 2.[1].54 different features are taken to classifying theses tri-dosha.

**Table 2**. Characteristics of Tri-Doshas [1].

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No. | Characteristics | KAPHA | PITTA | VATA |
| 1 | Body Frame | Large (thick skin) | Medium Frame | Thin Frame |
| 2 | Body Weight | Excess weight | Moderate weight | Thin and lean |
| 3 | Skin type | Thick, oily, cold | Soft, warm, moist | Dry and rough |
| 4 | Eyes | Big, large, wide, prominent. | Medium size | Sunken, small |
| 5 | Hair | Thick, dark, soft | Thin, fine, silky | Dry, scanty, coarse, curly |
| 6 | Complexion | Fair, bright | Yellowish, reddish | Brown |
| 7 | Physical traits | Slow Digestion, good stamina, sound sleepy. | Good Digestion, strong metabolism, body temperature higher | Disturbed Sleeping, cold hands, crave sweet, sour salty. |
| 8 | Miscellaneous traits | Relaxed, loving, faithful | Focused, competitive | Lively, fun, full of joy |

**B. TRI-DOSHA Prediction and Diet Suggestion:** Tri-Dosha analysis using computer vision is concerned with automating the physical feature analysis to gain the visual understanding of human body’s elemental humour constitution [6]. Recently Anjali Goyal and Vishu Madaan has worked upon predicting Ayurveda constitution by collecting data from questionaries and validated it by Ayurveda experts [1]. Same as author collected 54 characteristic data set as shown in table. 2. Various research and model has been developed for this problem “*TRI-DOSHA Prediction and Diet Suggestion”* but a quantitative and valid approach was missing. As Ayurveda is vast field of research and development. In our approach for this problem some machine learning algorithm trained on this dataset for predicting the dosha and after predicting the dosha the diet food and suggestions to balance the vikriti [6].

**C. Decision Tree Algorithm and Chatbot:** In our approach to analyse and to predict the *prakriti (dosha)* of human body 54 characteristics of tri-dosha has used table. 2. Various machine learning algorithm like Support Vector Machine (SVM), Naïve Bayes, K-Nearest Neighbour (KNN), Decision Tree has been trained on this data. decision fitted properly on training data and gives the best performance. Using chatbot model will ask questionnaires about characteristics of tri-dosha and then predict which dosha is pre dominant in body.

**II Related Work:** Ayurveda is vast field for research purpose. many researchers have researched in Ayurveda for better human lifestyle and nature. Anjali, Madhav *et al.* [1][8]. As the *Ayurveda is* oldest medical field but since the principles like tri-dosha and all are not much popular [3]. The authors discussed the current trend of Ayurveda that need to be known to all and its important role in health care. There is a discussion of the aspects that need to be considered for the promotion and development of Ayurveda medicine [1]. Anjali, Vishu has researched in this field as VATA, PITTA, KAPHA are combination of *panchbhutas* [1]. Gayatri garde in their research also described about Prakriti of human body and used many methods to detect prakriti. Rama Jayasundar et all [7], discussed that the ability of machine learning algorithms to solve complex tasks with dynamic environment has contributed to its success in their research. The work done so far in the field of Ayurveda use of machine learning in health care monitoring, health healing has motivated for human *prakriti* examination in human body. Preeti Verma et all [8]. discussed in their research that Agni apart from the digestive function it is also responsible for the production of strength which has two aspects namely strength to resist the occurrence of disease and decay in the human body [8]. These research works motivated a lot to research in this field (Ayurveda Tri-Dosha).

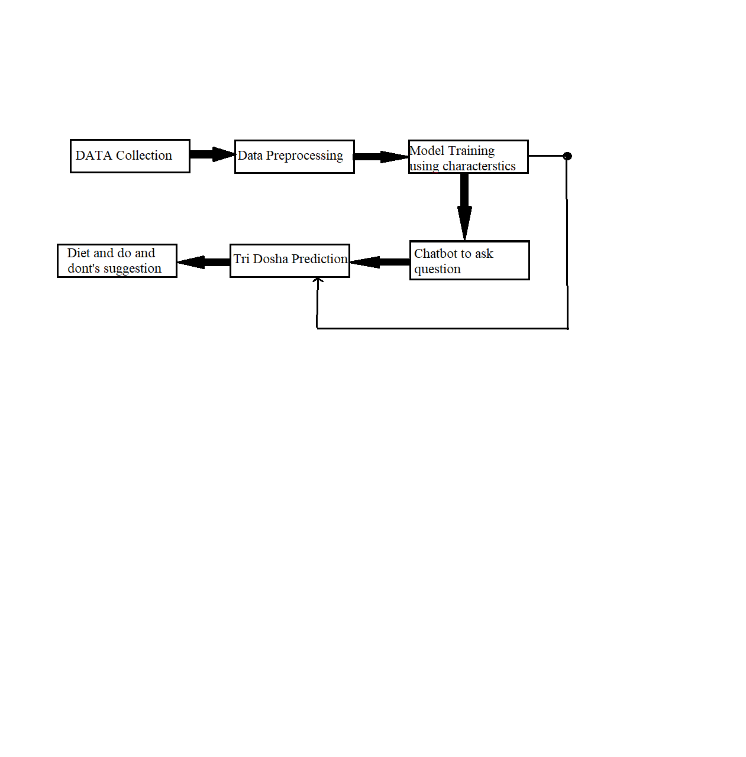
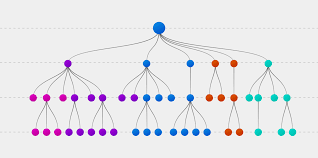
**III Methodology:** Here in this section, we have explained the methodology which is used in this research *Figure. 2*.

Figure.2 This is the work flow diagram fo achieving objectives after collecting data and Preprocessing ML model has trained with chatbot to predict the prkriti of dosha and according to prakriti the diet suggestions to balance dosha.

1. **Data Collection for Tri-Dosha Classification:**  In health care fields it is very rare that any hospital share there patients data in public and for this problem data is not on online platforms so by inspiring from Vishu ,Anjali[1] as they have collected the data of doshas characteristic by volunteering and validated it by ayurveda experts. Authors collected 54 characteristics of dataset through all prakriti (Tri-dosha) can easily classify *table.2*. After identifying the scheme and characteristics of all the different categories of doshas, we prepared a questionnaire for chatbot. the data which is prepared is easy to approach and less effort done to collect it.
2. **Data Validation:** The questionnaire prepared from the characteristics of doshas is validated from various online platform and by ayurveda experts to achieve better result and less error.
3. **Data Pre-processing:** the data used in this problem is collected by volunteering and by questionaries. So, the data is not balanced and biased also some methods are used to pre-process this data like *Encoding, outlier detection,* Splited (sampling) the data into 3:1 ratio to remove biasness.

1). Data Management: we have used Python modules and libraries like pandas, NumPy, sklearn for data pre-processing. The data is manually converted into sparse matrix from the``raw'' form using *Label Encoding* [1].

2). Imbalance Data: A proportionate collection of data collected from different classes is considered as best for learning of algorithm and good accuracy. Modest class imbalance can cause serious problems in training [1]. A number of different techniques are there to resolve this type of imbalanced class problem which focus either at the data level (sampling methods) or at the classifier level (modifying it internally).

3). Incomplete Data: while collecting the data there is chances of missing values as we have collected the data manually there were some missing values (nan values) it fills these values we have used imputation like simple imputer [1].

a. Imputation: it is technique to fill the missing data in the dataset it is the estimated value of the given information *mean, mode and median* is used to fill the missing data (nan value).

**IV Machine Learning Model:**

1. Decision Tree: The Decision Tree algorithm belongs to the deterministic family of classifiers and is a model-based technique [12]. The DT algo creates a tree like structure from the training dataset as used. This tree is then traversed to generate predictions. Despite all its flaws, it can work well with moderately large and complex datasets with 'n' different classes [1]. A decision tree consists of the top parent node as root, lower child nodes, branches and internal nodes, leaf node or also known as value node.in decision tree *information gain, entropy* and *Gini index* is used as thresh hold for a node.

1). Information Gain: To find the best feature which serves as a root node in terms of information gain, we first use each descriptive feature and split the dataset along the values of these descriptive features and then calculate the entropy of the dataset [1].

**2). Gini Index: It** is calculated by subtracting the sum of squared probabilities of each class from one. It favours larger partitions and easy to implement whereas information gain favours smaller partitions with distinct values [5].

Figure 3: Decision Tree Diagram.

Dataset is best fitted on decision tree model. model has archived 92% accuracy on Training dataset of doshas characteristics.

**V. Result and Analysis:** In our research we have found that there are several researchers have researched on this topic “*Tri-Dosha Prediction*” like Mohanpriya [14] et all. Yet these all models and research work was there but a proper systematic model is still missing that can predict the doshas as well as the diet and do and don’ts to balance the imbalanced *prakriti.* considering it we have trained ML model with dataset to predict the dosha and the diet plan to maintain the nature of prakriti and archived best performance on decision tree 0.95 precision,0.78 recall and 0.83 accuracy rate.

**VI. Conclusion:** The main goal of this research work is to predict the Tri-doshas in human body perfectly so that people’s can make balance in their lifestyle and nature [10]. Model has trained on various characteristics of eternal humour (prakriti). collected data by volunteering and validated it by online Ayurveda websites and experts and then train several ML Algorithm, Decision Tree fitted best on this data as the performance was high after training, we have again cross checked the predicted result manually to reduce error. after predicting the dosha successfully the model will predict the diet for that dosha also to balance that dosha. Every healthy or unhealthy individual can use this system without consulting an Ayurveda Expert to know about Dosha imbalances in one’s body and balance it directly by following suggested diet plan.

**VIII References:**

[1]. Akhil, C., Kumar, M. V. V., & Anoop. (2020).

An explorative conceptual study of samanya visesha and tridosha siddhanta based on systems biology: A review. *International Journal of Research in Ayurveda and Pharmacy*, *11*(2), 87–89.

[2]. Akhtar Roji, A. Y., & Paliwal, M. (2020). Concept of Dosha in Ayurveda: A review. *International Journal of Research in Ayurveda and Pharmacy*, *11*(6), 105–111.

[3]. Al-Shura, A. N. (2020). Tridosha and pathology. In *Perspectives of Ayurveda in Integrative Cardiovascular Chinese Medicine for Patient Compliance* (pp. 29–31). Elsevier.

[4]. Ayurvedic Body Type. (n.d.). Retrieved May 5,2021, from Dabur.com website: <https://www.dabur.com/in/enus/about/science-of-ayurveda/ayurvedic-body-types>

[4]. Bagali, D. S. S., & Baragi, U. C. (2016). Concept of Manasa Prakruti as described in Charaka Samhita. *Journal of Ayurveda and Integrated Medical Sciences (JAIMS)*, *1*(3). doi:10.21760/jaims. v1i3.4423

[5]. Bhargav, H., Jasti, N., More, P., Kumar, V., Chikkanna, U., Kishore Kumar, R., & Varambally, S. (2021). Correlation of prakriti diagnosis using AyuSoft prakriti diagnostic tool with clinician rating in patients with psychiatric disorders. *Journal of Ayurveda and Integrative Medicine*. doi: 10.1016/j.jaim.2021.01.012

[6]. Edavalath, M., & Bharathan, B. P. (2021). Methodology for developing and evaluating diagnostic tools in Ayurveda - a review. *Journal of Ayurveda and Integrative Medicine*. doi: 10.1016/j.jaim.2021.01.009

[7]. Gandhi, Z. (2020, May 4). What exactly are the 3 Ayurvedic doshas? Retrieved May 5, 2021, from Vedix.com

[8]. Hankey, A. (2005). A test of the systems analysis underlying the scientific theory of Ayurveda’s Tridosha. *Journal of Alternative and Complementary Medicine (New York, N.Y.)*, *11*(3), 385–390.

[9]. J, D. S., Sk, H., V., A. M., & Nalinakshan, A. (2018). A Survey on Menopausal Symptoms in relation to Prakruti. *Journal of Ayurveda and Integrated Medical Sciences (JAIMS)*, *3*(3). doi:10.21760/jaims. v3i3.12877

[10]. Joshi, R. R. (2004a). A biostatistical approach to ayurveda: quantifying the tridosha. *Journal of Alternative and Complementary Medicine (New York, N.Y.)*, *10*(5), 879–889.

[11]. Joshi, R. R. (2004b). A biostatistical approach to Ayurveda: Quantifying the tridosha. *Journal of Alternative and Complementary Medicine (New York, N.Y.)*, *10*(5), 879–889.

[12]. Lakhotia, S. C. (2014). Translating Ayurveda’s Dosha-Prakriti into objective parameters. *Journal of Ayurveda and Integrative Medicine*, *5*(3), 176.

[13]. Lokesh, T. (2016). Correlation of hypotension and hypertension with vata, pita and kapha as basic constituents. *Journal of Pharmaceutical and Scientific Innovation*, *5*(1), 7–11.

[14]. Mohanapriya, D. (2021). Predicting ayurvedha-based constituent balancing in human body using machine learning methods. *International Journal for Research in Applied Science and Engineering Technology*, *9*(3), 1413–1418.

[15]. Shukla, A. K., Mall, M., Rai, S. K., Singh, S., Nair, P., Parashar, G., … Khanuja, S. P. S. (2013). A transcriptomic approach for exploring the molecular basis for dosha-balancing property-based classification of plants in Ayurveda. *Molecular Biology Reports*, *40*(4), 3255–3262.

[16]. The Three Doshas in Ayurveda. (2020, October 19). Retrieved May 5, 2021, from Theayurveda.org.

[17]. Vinod, & Patil, U. (2020). A conceptual study on tridosha perspective of mild cognitive impairment: A review. *International Journal of Research in Ayurveda and Pharmacy*, *11*(5), 150–153.