FINAL YEAR PROJECT REPORT

Predicting Ayurveda-Based Constituent Balancing in Human Body Using

Machine Learning Methods

Bachelor of Technology in

Information Technology

Submitted by

Name of student	ID number
Anish Ambekar	171080061
Shreyash Vijaykar	171080075
Tushar Chaure	171080020
Charudatt Mahajan	171080021

Under the guidance of

Prof. V K Sambhe

Department of Computer Engineering and

Information Technology

VEERMATA JIJABAI TECHNOLOGICAL INSTITUTE

Mumbai , India – 400019

Semester VIII, 2021

Github Repository - https://github.com/TusharChaure/FYP
Project url - https://prakruti-prediction.herokuapp.com/

DECLARATION OF CANDIDATES

We state that work embodied in this Project entitled "Predicting Ayurveda-Based Constituent Balancing in Human Body Using Machine Learning Methods" forms our own contribution of work under the guidance Prof. V. K. Sambhe at the Department of Computer Engineering and Information Technology, Veermata Jijabai Technological Institute. The report reflects the work done during the period of candidature but may include related preliminary material provided that it has not contributed to an award of previous degree. No part of this work has been used by us for the requirement of another degree except where explicitly stated in the body of the text and the attached statement.

Signature of Candidates:

Tushar Chaure_ (171080020)

Shreyash Vijaykar_(171080075)

Anish Ambekar_ (171080061)

Charudatt Mahajan_ (181080021)

Date: 15/05/2021_____

ACKNOWLEDGEMENT

We would like to take this opportunity to express our sincere gratitude towards our mentor and project guide, Prof. V. K. Sambhe for his constant support and encouragement.

We would like to thank her for devoting her time and effort to this project. His words of encouragement and wisdom have kept us on track throughout this project and she has helped us achieve our results.

We would also like to take this opportunity to thank our parents for bearing with us and taking care of us while we set about this journey. They have been the pillars of support that helped us work on this project with dedication and spirit.

Finally, we would like to express our gratitude towards our alma mater, Veermata Jijabai

Technological Institute, for the extremely conducive environment it has provided and the infrastructure it has made available to students like us to pursue our projects also during the Covid Pandemic.

Table of contents :

Topic	Page no
Overview	5
Problem statement	6
Introduction	7
Methodology	11
Implementation Results and Performance analysis	20
UI Design	27
Conclusion	30
References	31

❑ Chapter 1: Overview

Human Body constitution (*prakriti*) defines what is in harmony with human nature and what

will cause to move out of balance and experience illness. *Tridosha* defines the three basic energies or principles that determine the function of our body on the physical and emotional levels.

The classification of people into the three fundamental constituents of Vata, Pitta, and Kaph is the foundation of every Ayurvedic treatment. This theory is called “Tri- dosha” in Ayurveda where ‘tri’ stands for three and ‘Dosha’ stands for constituents in Sanskrit language.

Tri-dosha are continuously formed in the human body and they are responsible for the homeostasis and metabolic activities in the human body. Tridosha influence the mind and body functions. The proportion of Tridosha constituents is unique for every person and is termed as his/her “Prakruti” where ‘pra’ means original and ‘kruti’ means creation in Sanskrit language. Prakruti determines the physiological and psychological tendencies in all living beings, influencing their physiological features like the texture and color of the skin, hair, eyes, length of fingers, the shape of the palm, body frame, strength of digestion etc., and psychological features like their nature (introverted, extroverted, calm, excitable, intense, laid- back), including their reaction to stress.

Prakruti usually does not change throughout the lifetime of a person and forms the basis behind studying, how different people respond to the same environmental conditions in varied ways. It also helps to assesses the patient’s suitability for a treatment.

Usually, only Ayurveda specialist or person who has knowledge of Ayurveda can predict the prakriti of any person based on his/her body constituents but to overcome this problem we are developing a website where anyone can take his/her prakriti analysis test to know their prakriti at free of cost in less than a min. So, for this we are using 5 different Machine Learning algorithms SVM, KNN, Decision Tree, Naïve Bayes and ANN at backend of website to analyze required prakriti of person based on given input parameters.

❑ Chapter 2: Problem Statement

Rather than indifferently prescribing the same treatments and chemicals to all people Ayurveda treats people individually according to their constitution and personality which can't be the same for any two persons. Every person is dominated by a unique combination of influencing energies (Tridosha Theory) which affect everyone slightly differently causing individual vulnerabilities and resistances.

So instead of chemically putting the fire out once it occurs Ayurveda focuses on preserving health by gently and naturally balancing energies and their influences to help people using their own self-healing powers.

❑ Chapter 3: Introduction

Ayurveda, a 5000-year-old science of healthcare or medical science, is the oldest science of healing known to the world. Ayurveda believes in the idea of embedding wellbeing into the human lifestyle to prevent and cure diseases. It focuses on strengthening and maintaining the balance in the body so that its natural defense system is strong enough to protect against diseases. Just as every person has a unique fingerprint, every person also has unique physical, mental and emotional characteristics. These characteristics originate from a unique proportion of three principal constituents, which is determined at the conception of the living being and remain fixed throughout one's lifetime. Human nature (prakriti) has two variants: the physical(Sharirika Prakriti) and the intellectual (Manasic Prakriti). Vata, pitta and kaph are Sharirika prakriti

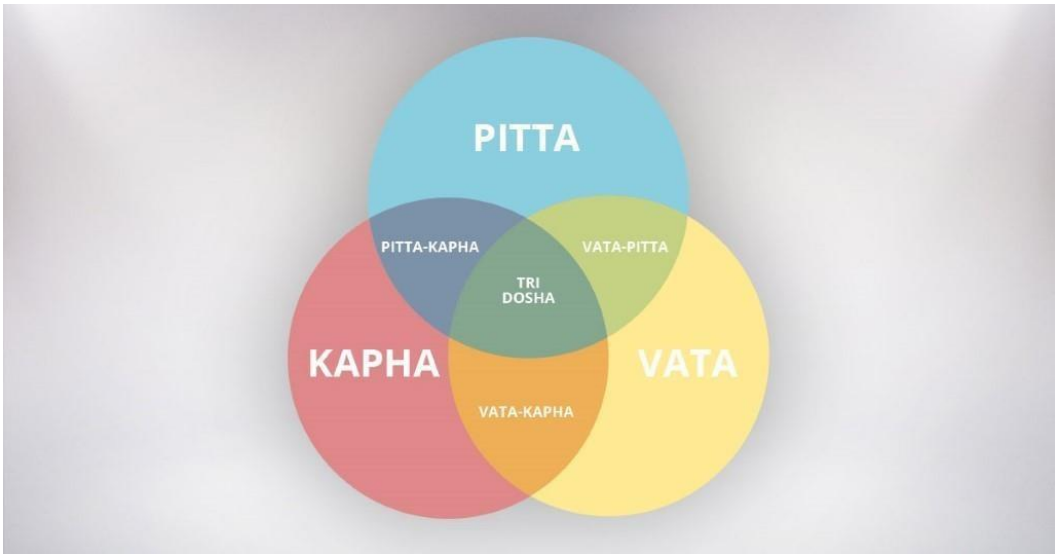
The three principal constituents are Vata, Pitta and Kaph. Ayurveda believes that every element in the universe is made up of five building blocks namely, air, fire, water, space and earth. Vata, Pitta and Kaph originated from the combination of these elements.

Vata is the energy of movement and is composed of air and space. It governs breathing, pulsation of heart, blinking, muscle and tissue movement and all the movements in the cytoplasm and cell membranes. Vata promotes creativity and flexibility when in balance and fear and anxiety when out of balance. Vata people are more prone to diseases associated with the air principle, like nervous system related disorders, arthritis, emphysema, Alzheimer, etc. Pitta is the energy of digestion and metabolism which is composed of fire and water elements. It is responsible for digestion, metabolism, absorption, assimilation and the body temperature.

Pitta promotes understanding and intelligence when in balance and, anger and jealousy when out of balance.

Kaph is the energy of lubrication and structure which is made up of earth and water. It is responsible for forming the structure of the body by holding the cells together via bones, tendons and muscles. Kaph promotes love, calmness and forgiveness when in balance and envy and greed when out of balance.

Following table shows 7 different types in which Prakruti can be classified. Most people will have one predominant Dosha Vatt, PITT, KAFF (V, P or K), comparatively few people have two Dosha equal in proportion VATT-PITT, VATT- KAFF, PITT-KAFF (VP, VK, PK) and even fewer will have all the three Doshas in equal proportion VATT-PITT-KAFF(VPK).



Types of Prakruti :

TYPE	DESCRIPTION
V, P, K	Predominant in one Dosha
VP, VK, PV	Two relatively equal proportions with one predominating
VPK	Doshas in almost equal proportion

Dominant Dosha for Prakruti analysis can be discovered by observing physiological and psychological features of a person. Following table enlists few special characteristics observed by Ayurvedic doctors to determine the Prakruti of people.

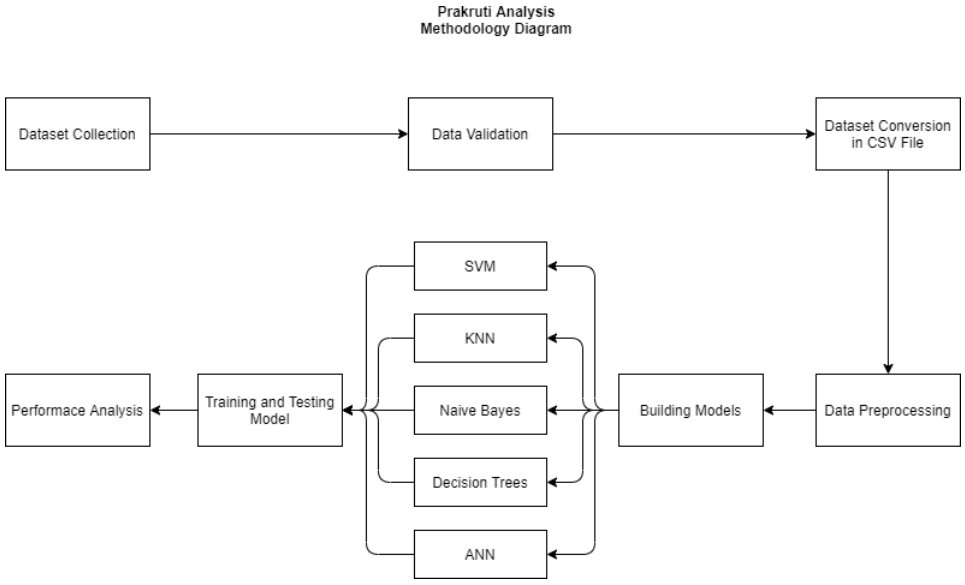
Knowing the dosha type beforehand helps in planning the lifestyle and diet according to the body's needs. This knowledge provides probable occurrence of qualitative and quantitative imbalances in the body

NO	OBSERVATION	TRIDOSHA		
		VATT	PITT	KAFF
1	Body size	Slim	Medium	Large
2	Body weight	Low	Medium	Overweight
3	Cheeks	Wrinkled	Smooth flat	Rounded, plump
4	Face shape	Thin, angular	Tapering/triangular	Rounded, double chin
5	Eyes	Small, sunken, dry, active, black, brown, nervous	Sharp, bright, gray, green, yellow/red, sensitive to light	Big, beautiful, blue, calm, loving
6	Nose	Uneven, deviated septum	Long pointed, red nose-tip	Short rounded, button nose
7	Lips	Dry, cracked, black/brown tinge	Red, inflamed, yellowish	Smooth, oily, cool, white, pale
8	Teeth	Stick out, big, roomy, thin gums	Medium, soft, tender gums	Healthy, white, strong gums
9	Skin	Thin, dry, cold, rough, dark	Smooth, oily, warm, rosy	Thick, oily, cool, white, pale
10	Hair	Irregular, scanty	Strong, unbearable	Slow but steady
11	Appetite	Irregular, forms gas	Quick, causes burning	Prolonged, forms mucous
12	Height	Tall or Short	Average	Thin and sturdy / Short and Stocky
13	Bone Structure	Light, small bones, prominent joints	Medium bone structure	Large, Broad shoulders, Heavy bone structure
14	Complexion	Dark complexion, tans easily	Fair skin, sun burns easily	White, pale, tans evenly
15	Texture of skin	Dry, pigmentation and aging	Freckles, many moles, redness, rashes and acne	Soft, glowing and youthful
16	Hair color	Dull, black, brown	Red, light brown, yellow	Brown

17	Appearance of hair	Dry, Black knotted, Brittle	Straight, Oily	Thick, Curly
18	Eyelashes	Scanty eye lashes	Moderate eye lashes	Thick / Fused eye lashes
19	Blinking of eye	Excessive blinking	Moderate blinking	More or less stable
20	Cheeks	Wrinkled, Sunken	Smooth, flat	Rounded, plump

❑ Chapter 4 : Methodology and Implementation

This Chapter explains the whole process of project starting from dataset collection to different implementation techniques.



Step 1 : Dataset Collection

As dataset for given topic was not available on google, so we create our own dataset with the help of survey. In this survey we created one google form with 20 different question based different body constituents' like hair color, body weight, Skin color, etc. as mentioned in above introduction part to analyze Prakruti of person using different Machine Learning models.

Google Form link :

https://docs.google.com/forms/d/e/1FAIpQLSdZGKoC0iqe4oq4odN9STGvw820kQ8pX7KhOXgdoVtUkSPbxA/viewform?usp=sf_link

Step 2 : Data Validation

After one and half month we received around 1200 records from our survey which were not analyze yet. For any machine learning model good dataset is necessary hence to analyze results of given records we take help of Ayurveda specialist who first manually analyze first 25 records and she also confirmed each result on prakriti analysis site and she also recommend us to use same site to analyze result of each record. So for around 1150 remaining results prediction we use same site and make one more entry column in excel sheet for results and named it as Prakruti.

As whole dataset was in text format, so we decided to modify each record in numerical format. Using find and replace option we convert whole records in numerical format where independent attributes converted in range of 0 to 2 and dependent variable into 0 to 5 range as prakriti is classified into 6 different types (For simplicity purpose we eliminated all records of V-P-K prakriti).

Here, for Independent variables

0 Indicates VATT(V)

1 Indicates PITT(P)

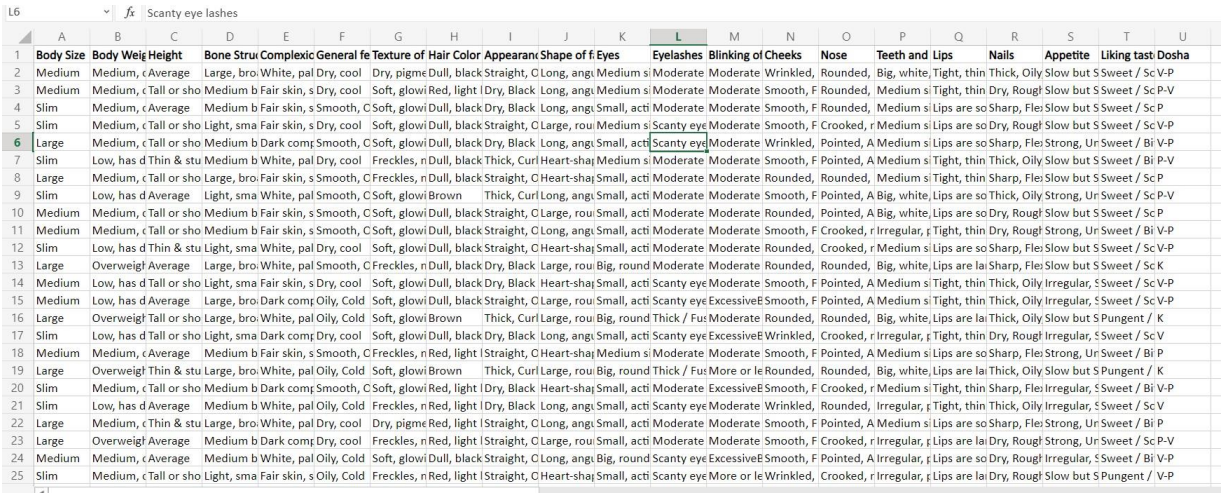
2 Indicates KAFF(K)

3 Indicates VATT-PITT(V-P)

4 Indicates PITT-KAFF(P-K)

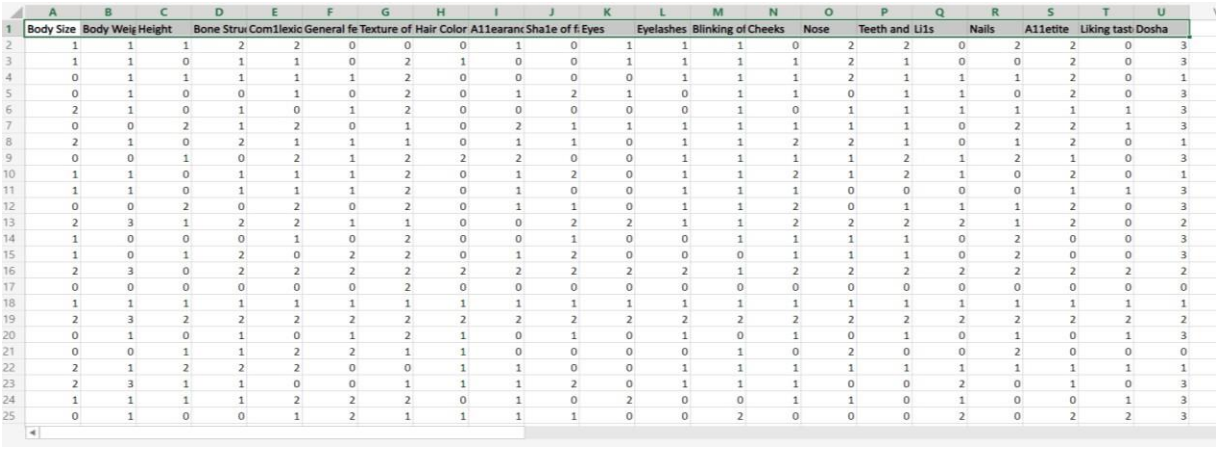
5 Indicates VATT-KAFF(V-K)

Dataset in texted format :



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	Body Size	Body Weig	Height	Bone Stru	Complexic	General fe	Texture of	Hair Color	Appearanc	Shape of f	Eyes	Eyelashes	Blinking of	Cheeks	Nose	Teeth and	Lips	Nails	Appetite	Liking tast	Dosha
2	Medium	Medium, (	Average	Large, bro	White, pal	Dry, cool	Dry, pigme	Dull, black	Straight, C	Long, angu	Medium si	Moderate	Moderate	Wrinkled,	Rounded,	Big, white	Tight, thin	Thick, Oily	Slow but S	Sweet / Sc	V-P
3	Medium	Medium, (	Tall or sho	Medium b	Fair skin, s	Dry, cool	Soft, glowi	Red, light I	Dry, Black	Long, angu	Medium si	Moderate	Moderate	Smooth, F	Rounded,	Medium si	Tight, thin	Dry, Rough	Slow but S	Sweet / Sc	P-V
4	Slim	Medium, (	Average	Medium b	Fair skin, s	Smooth, C	Soft, glowi	Dull, black	Dry, Black	Long, angu	Small, acti	Moderate	Moderate	Smooth, F	Rounded,	Medium si	Lips are so	Sharp, Fle	Slow but S	Sweet / Sc	P
5	Slim	Medium, (	Tall or sho	Light, sma	Fair skin, s	Dry, cool	Soft, glowi	Dull, black	Straight, C	Large, rou	Medium si	Scanty eye	Moderate	Smooth, F	Crooked, r	Medium si	Lips are so	Dry, Rough	Slow but S	Sweet / Sc	V-P
6	Large	Medium, (	Tall or sho	Medium b	Dark com	Smooth, C	Soft, glowi	Dull, black	Dry, Black	Long, angu	Small, acti	Scanty eye	Moderate	Wrinkled,	Pointed, A	Medium si	Lips are so	Sharp, Fle	Strong, Ur	Sweet / Bi	V-P
7	Slim	Low, has d	Thin & stu	Medium b	White, pal	Dry, cool	Freckles, n	Dull, black	Thick, Curl	Heart-shar	Medium si	Moderate	Moderate	Smooth, F	Pointed, A	Medium si	Tight, thin	Thick, Oily	Slow but S	Sweet / Bi	P-V
8	Large	Medium, (	Tall or sho	Large, bro	Fair skin, s	Smooth, C	Freckles, n	Dull, black	Straight, C	Heart-shar	Small, acti	Moderate	Moderate	Rounded,	Rounded,	Medium si	Tight, thin	Sharp, Fle	Slow but S	Sweet / Sc	P
9	Slim	Low, has d	Average	Light, sma	White, pal	Smooth, C	Soft, glowi	Brown	Thick, Curl	Long, angu	Small, acti	Moderate	Moderate	Smooth, F	Pointed, A	Big, white	Lips are so	Thick, Oily	Strong, Ur	Sweet / Sc	P-V
10	Medium	Medium, (	Tall or sho	Medium b	Fair skin, s	Smooth, C	Soft, glowi	Dull, black	Straight, C	Large, rou	Small, acti	Moderate	Moderate	Rounded,	Pointed, A	Big, white	Lips are so	Dry, Rough	Slow but S	Sweet / Sc	P
11	Medium	Medium, (	Tall or sho	Medium b	Fair skin, s	Smooth, C	Soft, glowi	Dull, black	Straight, C	Long, angu	Small, acti	Moderate	Moderate	Smooth, F	Crooked, r	Irregular, (	Tight, thin	Dry, Rough	Strong, Ur	Sweet / Bi	V-P
12	Slim	Low, has d	Thin & stu	Light, sma	White, pal	Dry, cool	Soft, glowi	Dull, black	Straight, C	Heart-shar	Small, acti	Moderate	Moderate	Rounded,	Crooked, r	Medium si	Lips are so	Sharp, Fle	Slow but S	Sweet / Sc	V-P
13	Large	Overweigh	Average	Large, bro	White, pal	Smooth, C	Freckles, n	Dull, black	Dry, Black	Large, rou	Big, round	Moderate	Moderate	Rounded,	Rounded,	Big, white	Lips are lai	Sharp, Fle	Slow but S	Sweet / Sc	K
14	Medium	Low, has d	Tall or sho	Light, sma	Fair skin, s	Dry, cool	Soft, glowi	Dull, black	Dry, Black	Heart-shar	Small, acti	Scanty eye	Moderate	Smooth, F	Pointed, A	Medium si	Tight, thin	Thick, Oily	Irregular, S	Sweet / Sc	V-P
15	Medium	Low, has d	Average	Large, bro	Dark com	Oily, Cold	Soft, glowi	Dull, black	Straight, C	Large, rou	Small, acti	Scanty eye	ExcessiveE	Smooth, F	Pointed, A	Medium si	Tight, thin	Thick, Oily	Irregular, S	Sweet / Sc	V-P
16	Large	Overweigh	Tall or sho	Large, bro	White, pal	Oily, Cold	Soft, glowi	Brown	Thick, Curl	Large, rou	Big, round	Thick / Fus	Moderate	Rounded,	Rounded,	Big, white	Lips are lai	Thick, Oily	Slow but S	Pungent /	K
17	Slim	Low, has d	Tall or sho	Light, sma	Dark com	Dry, cool	Soft, glowi	Dull, black	Dry, Black	Long, angu	Small, acti	Scanty eye	ExcessiveE	Wrinkled,	Crooked, r	Irregular, (	Tight, thin	Dry, Rough	Irregular, S	Sweet / Sc	V
18	Medium	Medium, (	Average	Medium b	Fair skin, s	Smooth, C	Freckles, n	Red, light I	Straight, C	Heart-shar	Medium si	Moderate	Moderate	Smooth, F	Pointed, A	Medium si	Lips are so	Sharp, Fle	Strong, Ur	Sweet / Bi	P
19	Large	Overweigh	Thin & stu	Large, bro	White, pal	Oily, Cold	Soft, glowi	Brown	Thick, Curl	Large, rou	Big, round	Thick / Fus	More or le	Rounded,	Rounded,	Big, white	Lips are lai	Thick, Oily	Slow but S	Pungent /	K
20	Slim	Medium, (	Tall or sho	Medium b	Dark com	Smooth, C	Soft, glowi	Red, light I	Dry, Black	Heart-shar	Small, acti	Moderate	ExcessiveE	Smooth, F	Crooked, r	Medium si	Tight, thin	Sharp, Fle	Irregular, S	Sweet / Bi	V-P
21	Slim	Low, has d	Average	Medium b	White, pal	Oily, Cold	Freckles, n	Red, light I	Dry, Black	Long, angu	Small, acti	Scanty eye	Moderate	Wrinkled,	Rounded,	Irregular, (	Tight, thin	Thick, Oily	Irregular, S	Sweet / Sc	V
22	Large	Medium, (	Thin & stu	Large, bro	White, pal	Dry, cool	Dry, pigme	Red, light I	Straight, C	Long, angu	Small, acti	Moderate	Moderate	Smooth, F	Pointed, A	Medium si	Lips are so	Sharp, Fle	Strong, Ur	Sweet / Bi	P
23	Large	Overweigh	Average	Medium b	Dark com	Dry, cool	Freckles, n	Red, light I	Straight, C	Large, rou	Small, acti	Moderate	Moderate	Smooth, F	Crooked, r	Irregular, (	Lips are lai	Dry, Rough	Strong, Ur	Sweet / Sc	P-V
24	Medium	Medium, (	Average	Medium b	White, pal	Oily, Cold	Soft, glowi	Dull, black	Straight, C	Long, angu	Big, round	Scanty eye	ExcessiveE	Smooth, F	Pointed, A	Irregular, (	Lips are so	Dry, Rough	Irregular, S	Sweet / Bi	V-P
25	Slim	Medium, (	Tall or sho	Light, sma	Fair skin, s	Oily, Cold	Freckles, n	Red, light I	Straight, C	Heart-shar	Small, acti	Scanty eye	More or le	Wrinkled,	Crooked, r	Irregular, (	Lips are lai	Dry, Rough	Slow but S	Pungent /	V-P

Dataset in Numerical Format :



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	Body Size	Body Weig	Height	Bone Stru	Com1lexic	General fe	Texture of	Hair Color	A11earanc	Sha1e of f	Eyes	Eyelashes	Blinking of	Cheeks	Nose	Teeth and	Li1s	Nails	A11etite	Liking tast	Dosha
2	1	1	1	2	2	0	0	0	1	0	1	1	1	0	2	2	0	2	2	0	3
3	1	1	0	1	1	0	2	1	0	0	1	1	1	1	2	1	0	0	2	0	3
4	0	1	1	1	1	1	2	0	0	0	0	1	1	1	2	1	1	1	2	0	1
5	0	1	0	0	1	0	2	0	1	2	1	0	1	1	0	1	1	0	2	0	3
6	2	1	0	1	0	1	2	0	0	0	0	0	1	0	1	1	1	1	1	1	3
7	0	0	2	1	2	0	1	0	2	1	1	1	1	1	1	1	0	2	2	1	3
8	2	1	0	2	1	1	1	0	1	1	0	1	1	2	2	1	0	1	2	0	1
9	0	0	1	0	2	1	2	2	2	0	0	1	1	1	1	2	1	2	1	0	3
10	1	1	0	1	1	1	2	0	1	2	0	1	1	2	1	2	1	0	2	0	1
11	1	1	0	1	1	1	2	0	1	0	0	1	1	1	0	0	0	0	1	1	3
12	0	0	2	0	2	0	2	0	1	1	0	1	1	2	0	1	1	1	2	0	3
13	2	3	1	2	2	1	1	0	0	2	2	1	1	2	2	2	2	1	2	0	2
14	1	0	0	0	1	0	2	0	0	1	0	0	1	1	1	1	0	2	0	0	3
15	1	0	1	2	0	2	2	0	1	2	0	0	0	1	1	1	0	2	0	0	3
16	2	3	0	2	2	2	2	2	2	2	2	2	1	2	2	2	2	2	2	2	2
17	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
19	2	3	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
20	0	1	0	1	0	1	2	1	0	1	0	1	0	1	0	1	0	1	0	1	3
21	0	0	1	1	2	2	1	1	0	0	0	0	1	0	2	0	0	2	0	0	0
22	2	1	2	2	2	0	0	1	1	0	0	1	1	1	1	1	1	1	1	1	1
23	2	3	1	1	0	0	1	1	1	2	0	1	1	1	0	0	2	0	1	0	3
24	1	1	1	1	2	2	2	0	1	0	2	0	0	1	1	0	1	0	0	1	3
25	0	1	0	0	1	2	1	1	1	1	0	0	2	0	0	0	2	0	2	2	3

Step 2 : Data Pre-processing

Once our dataset is ready then we will start our implementation part in which **Data Pre-processing** is the first step. We have to prepare some stuff in order to make sure that we build our machine learning models without any issues and this stuff that we need to prepare for machine learning trip happens to be data pre- processing.

For Data Pre-processing we will follow following steps :

1. Acquire the dataset:

Please refer data collection point.

2. Import all the crucial libraries:

Initially we will import essential libraries like Pandas, NumPy, Mat plot, sklearn, etc. Which are used for different activities like read data to perform different operations on dataset, etc.

3. Import the dataset:

Next step in data pre-processing is to read dataset from local storage for which we will use panda's library which we already imported in above step.

4. Identifying and handling the missing values:

Now we will start preparing the data so that our machine learning models run correctly and the first problem that we have to deal with is the case where you have some missing data in your data set and that happens quite a lot actually in real life. So, you have to get the trick to handle this problem and make it all good for your machine learning model to run correctly.

For this using inbuild libraries we will check is there any missing values present in our dataset or not if it's there then we will remove that record as we want better accuracy for our models. One more reason behind removing null values records is we can't take mean value in place of missing value as our dataset is purely in text format.

5. Encoding the categorical data:

For any machine learning model input as well as result values must be in numerical format. But in our dataset, we can see that we have 21 categorical variables as

shown in below screenshot where first line indicates all categorical variables names.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1	Body Size	Body Weig	Height	Bone Struc	Complexic	General fe	Texture of	Hair Color	Appearanc	Shape of f	Eyes	Eyelashes	Blinking of	Cheeks	Nose	Teeth and	Lips	Nails	Appetite	Liking tast	Dosha		
2	Medium	Medium, (	Average	Large, bro	White, pal	Dry, cool	Dry, pigme	Dull, black	Straight, C	Long, angu	Medium si	Moderate	Moderate	Wrinkled,	Rounded,	Big, white,	Tight, thin	Thick, Oily	Slow but S	Sweet / Sc	V-P		
3	Medium	Medium, (	Tall or sho	Medium b	Fair skin, s	Dry, cool	Soft, glowi	Red, light I	Dry, Black	Long, angu	Medium si	Moderate	Moderate	Smooth, F	Rounded,	Medium si	Tight, thin	Dry, Rough	Slow but S	Sweet / Sc	P-V		
4	Slim	Medium, (	Average	Medium b	Fair skin, s	Smooth, C	Soft, glowi	Dull, black	Dry, Black	Long, angu	Small, acti	Moderate	Moderate	Smooth, F	Rounded,	Medium si	Lips are so	Sharp, Fle	Slow but S	Sweet / Sc	P		
5	Slim	Medium, (	Tall or sho	Light, sma	Fair skin, s	Dry, cool	Soft, glowi	Dull, black	Straight, C	Large, rou	Medium si	Scanty eye	Moderate	Smooth, F	Crooked, r	Medium si	Lips are so	Dry, Rough	Slow but S	Sweet / Sc	V-P		
6	Large	Medium, (	Tall or sho	Medium b	Dark comp	Smooth, C	Soft, glowi	Dull, black	Dry, Black	Long, angu	Small, acti	Scanty eye	Moderate	Wrinkled,	Pointed, A	Medium si	Lips are so	Sharp, Fle	Strong, Ur	Sweet / Bi	V-P		
7	Slim	Low, has d	Thin & stu	Medium b	White, pal	Dry, cool	Freckles, n	Dull, black	Thick, Curl	Heart-shai	Medium si	Moderate	Moderate	Smooth, F	Pointed, A	Medium si	Tight, thin	Thick, Oily	Slow but S	Sweet / Bi	P-V		
8	Large	Medium, (	Tall or sho	Large, bro	Fair skin, s	Smooth, C	Freckles, n	Dull, black	Straight, C	Heart-shai	Small, acti	Moderate	Moderate	Rounded,	Rounded,	Medium si	Tight, thin	Sharp, Fle	Slow but S	Sweet / Sc	P		
9	Slim	Low, has d	Average	Light, sma	White, pal	Smooth, C	Soft, glowi	Brown	Thick, Curl	Long, angu	Small, acti	Moderate	Moderate	Smooth, F	Pointed, A	Big, white,	Lips are so	Thick, Oily	Strong, Ur	Sweet / Sc	P-V		
10	Medium	Medium, (	Tall or sho	Medium b	Fair skin, s	Smooth, C	Soft, glowi	Dull, black	Straight, C	Large, rou	Small, acti	Moderate	Moderate	Rounded,	Pointed, A	Big, white,	Lips are so	Dry, Rough	Slow but S	Sweet / Sc	P		
11	Medium	Medium, (	Tall or sho	Medium b	Fair skin, s	Smooth, C	Soft, glowi	Dull, black	Straight, C	Long, angu	Small, acti	Moderate	Moderate	Smooth, F	Crooked, r	Irregular, p	Tight, thin	Dry, Rough	Strong, Ur	Sweet / Bi	V-P		
12	Slim	Low, has d	Thin & stu	Light, sma	White, pal	Dry, cool	Soft, glowi	Dull, black	Straight, C	Heart-shai	Small, acti	Moderate	Moderate	Rounded,	Crooked, r	Medium si	Lips are so	Sharp, Fle	Slow but S	Sweet / Sc	V-P		
13	Large	Overweigh	Average	Large, bro	White, pal	Smooth, C	Freckles, n	Dull, black	Dry, Black	Large, rou	Big, round	Moderate	Moderate	Rounded,	Rounded,	Big, white,	Lips are la	Sharp, Fle	Slow but S	Sweet / Sc	K		
14	Medium	Low, has d	Tall or sho	Light, sma	Fair skin, s	Dry, cool	Soft, glowi	Dull, black	Dry, Black	Heart-shai	Small, acti	Scanty eye	Moderate	Smooth, F	Pointed, A	Medium si	Tight, thin	Thick, Oily	Irregular, S	Sweet / Sc	V-P		
15	Medium	Low, has d	Average	Large, bro	Dark comp	Oily, Cold	Soft, glowi	Dull, black	Straight, C	Large, rou	Small, acti	Scanty eye	Excessive E	Smooth, F	Pointed, A	Medium si	Tight, thin	Thick, Oily	Irregular, S	Sweet / Sc	V-P		
16	Large	Overweigh	Tall or sho	Large, bro	White, pal	Oily, Cold	Soft, glowi	Brown	Thick, Curl	Large, rou	Big, round	Thick / Fus	Moderate	Rounded,	Rounded,	Big, white,	Lips are la	Thick, Oily	Slow but S	Pungent /	K		
17	Slim	Low, has d	Tall or sho	Light, sma	Dark comp	Dry, cool	Soft, glowi	Dull, black	Dry, Black	Long, angu	Small, acti	Scanty eye	Excessive E	Wrinkled,	Crooked, r	Irregular, p	Tight, thin	Dry, Rough	Irregular, S	Sweet / Sc	V		
18	Medium	Medium, (	Average	Medium b	Fair skin, s	Smooth, C	Freckles, n	Red, light I	Straight, C	Heart-shai	Medium si	Moderate	Moderate	Smooth, F	Pointed, A	Medium si	Lips are so	Sharp, Fle	Strong, Ur	Sweet / Bi	P		
19	Large	Overweigh	Thin & stu	Large, bro	White, pal	Oily, Cold	Soft, glowi	Brown	Thick, Curl	Large, rou	Big, round	Thick / Fus	More or le	Rounded,	Rounded,	Big, white,	Lips are la	Thick, Oily	Slow but S	Pungent /	K		
20	Slim	Medium, (	Tall or sho	Medium b	Dark comp	Smooth, C	Soft, glowi	Red, light I	Dry, Black	Heart-shai	Small, acti	Moderate	Excessive E	Smooth, F	Crooked, r	Medium si	Tight, thin	Sharp, Fle	Irregular, S	Sweet / Bi	V-P		
21	Slim	Low, has d	Average	Medium b	White, pal	Oily, Cold	Freckles, n	Red, light I	Dry, Black	Long, angu	Small, acti	Scanty eye	Moderate	Wrinkled,	Rounded,	Irregular, p	Tight, thin	Thick, Oily	Irregular, S	Sweet / Sc	V		
22	Large	Medium, (	Thin & stu	Large, bro	White, pal	Dry, cool	Dry, pigme	Red, light I	Straight, C	Long, angu	Small, acti	Moderate	Moderate	Smooth, F	Pointed, A	Medium si	Lips are so	Sharp, Fle	Strong, Ur	Sweet / Bi	P		
23	Large	Overweigh	Average	Medium b	Dark comp	Dry, cool	Freckles, n	Red, light I	Straight, C	Large, rou	Small, acti	Moderate	Moderate	Smooth, F	Crooked, r	Irregular, p	Lips are la	Dry, Rough	Strong, Ur	Sweet / Sc	P-V		
24	Medium	Medium, (	Average	Medium b	White, pal	Oily, Cold	Soft, glowi	Dull, black	Straight, C	Long, angu	Big, round	Scanty eye	Excessive E	Smooth, F	Pointed, A	Irregular, p	Lips are so	Dry, Rough	Irregular, S	Sweet / Bi	V-P		
25	Slim	Medium, (	Tall or sho	Light, sma	Fair skin, s	Oily, Cold	Freckles, n	Red, light I	Straight, C	Heart-shai	Small, acti	Scanty eye	More or le	Wrinkled,	Crooked, r	Irregular, r	Lips are la	Dry, Rough	Slow but S	Pungent /	V-P		

And now you can guess that since machine learning models are based on mathematical equations you can intuitively understand that it would cause some problem if we keep the text here and the categorical variables in the equations because we would only want numbers in the equations. So that's why we need to encode the categorical variables. That is to encode the text that we have here into numbers. For this we will use sklearn inbuilt python libraries to encode all categorical variables in numbers.

6. Splitting the dataset:

Once label encoding is done then we need to divide our data set into two matrices. One for training(X) and other for testing(Y). According research paper we split our dataset in 80:20 ratio where around 800 records for training purpose and remaining 400 for testing. For splitting we will use sklearn inbuild python library.

7. Feature scaling:

Feature scaling is last step of Data Pre-processing. But for our dataset feature scaling is not needed as we have generated our own dataset for which all independent variables taken are essentials.

Hence we will skip this last step in implementation.

Step 3: Building Models

Third step of implementation is building machine learning model which will predict Prakriti of person for our project. As name of our project suggest that we have to

predict six different types of Prakrit's it clears that it’s a classification problem as our independent variables are of six different types which are as follows:

VATT, PITT, KAFF, VATT-PITT, PITT-KAFF, VATT-KAFF.

So, for this project we will build different classification models. For better understanding we will build all following five models given in research paper.

1. Naïve Bayes

Naive-Bayes is a classifier based on supervised learning or a statistical approach. Bayes' theorem-based approach assumes that the presence of a feature in a class is completely independent of the presence of any other feature. The Naive Bayes consider problem instances as feature vectors, which classified through the method to specific classes. These characteristics are not interdependent, that is, the value assigned to one characteristic does not impudence the value of other characteristics. Following Equation shows the relationship between the posterior probability P(T|A), P (T Class Prior Probability), P (A Predictor Prior Probability) and the probability P(A|T).

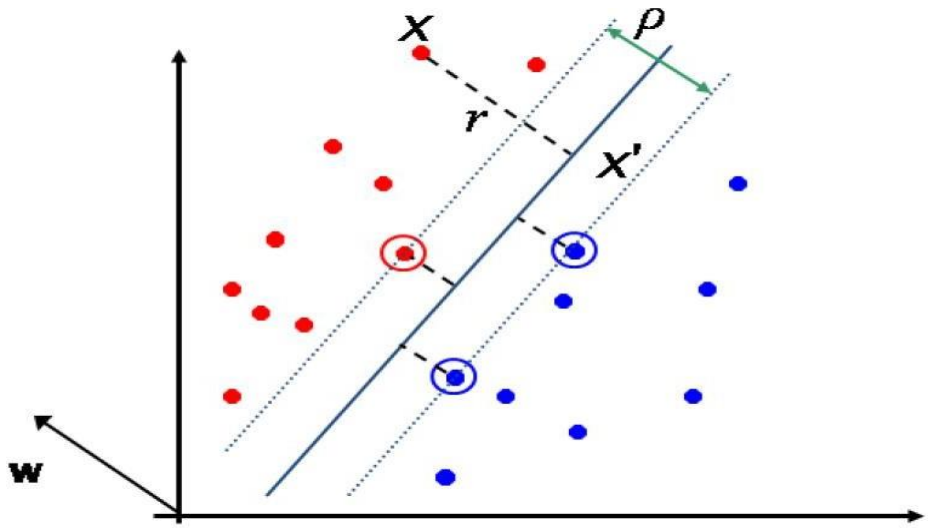
$$p(\frac{T}{A}) = \frac{p(\frac{A}{T}) * p(T)}{p(A)}$$

Where, T- Goal, A- Attributes.

2. SVM

SVM is a discriminatory classifier formally defined by a separating hyperplane. This hyperplane is a line that divides a plane into two parts, which in each class lie on either side in two-dimensional space. It is implemented on predefined, labeled training data with a monitored learning method. The algorithm outputs an optimal hyperplane that categorizes the novel data points. The algorithm of the support vector machine is supposed to find a hyperplane in an N-dimensional space that uniquely classifies the data points. Following figure shows the two different types of data points are separated by hyperplane. Distance form data points to separator is r as shown in following figure, is margin(width) between separator of classes as shown in figure. The objective is to maximizes the distance between the hyperplane and the difficult points close to decision boundary.

The reason behind high selection of SVM is because of its efficient implementations and performances proved to be excellent for high dimensional problems and small data sets.

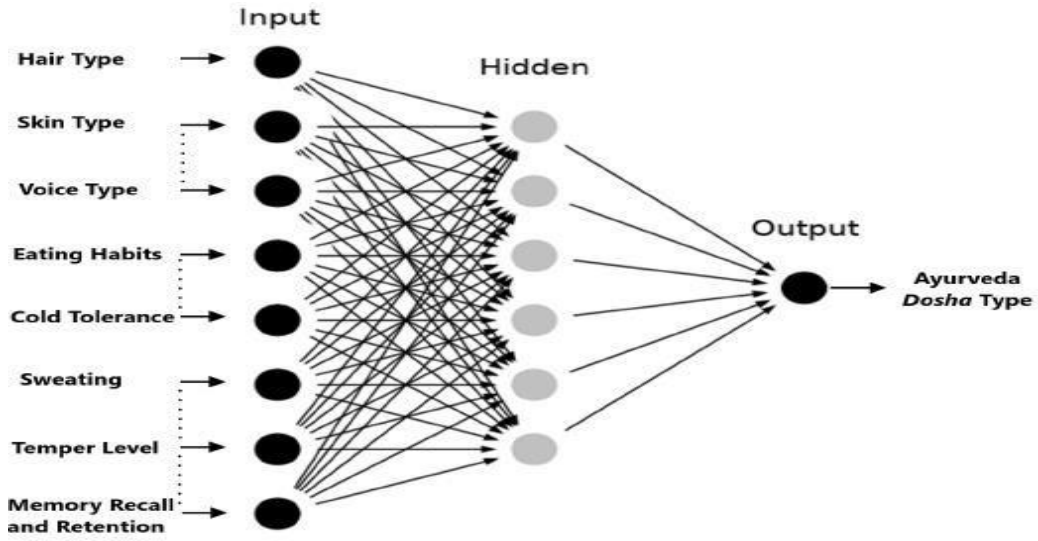


3. Decision Trees

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. The root connects the different classes of a tree. Leaf nodes represent the classes, the branches the results and the inner leaves the processes. The rules of classification rules form the paths from the root to the leaves. After a decision tree has been built from the data with the attributes A1, A2, . . . , An and the classes C1, C2, . . . , Cz, this decision tree can be used to classify a new data element with the attributes Aa, . . ., An into one of the classes C1, C2, . . . , Cz. For each new data element to be classified, each node including the root is considered a question for the data sample. Based on the response suggested by any available branch, the next node is selected. When accessing the next node, another question about the data sample is answered until it reaches the leaf node. A leaf node is connected to one of the classes C1, C2, . . . , Ck. In this way, we assign a certain class to the data sample.

4. ANN

The artificial neural network is purely similar to the human neural system. It consists of a number of artificial neurons that are trained to output the classification. These neurons are connected to weights that are adjusted to modulate the effect of the input signals. Significant input data are fed into the network with their output. Following figure shows what a coherent network looks like and connectivity of neurons on different layers.



For a defined training data set we select a suitable training model. Once the network is trained, we test ANN for its generalization performance. ANN gives feedback on whether the network has successfully classified the data or not. Many researchers use it ANN in medically related applications [30], [31]. Two types of activation functions are used in artificial neural networks:

I. Rectified Linear Unit (ReLU)

After the inputs from each layer have been passed, it is important to apply the activation function. ReLU is relatively simple, very useful and dynamic activation function. It offers the robustness to every small change like noise in the input. ReLU curve is half rectified. It means that for all negative input values, it turns the value into zero immediately.

I. SoftMax

We use the SoftMax activation function when the data is to be divided

into several classes. This function calculates the

probability distribution of k output classes. We will implement ANN in Python with Keras and TensorFlow libraries.

The ANN training parameters are listed in following table -

ANN Parameters	Values
Input layer Neurons	20
Activation Function	ReLU, SoftMax
Batch Size	100 / 128
Training Testing ratio	8:2

5. KNN

KNN algorithm is a supervised machine learning algorithm used for solving classification and regression problems. This is based on the fact that data elements assigned the same class if these are closer in a feature space. This technique is based on Euclidean distance method by which we can calculate the distance between two points in a coordinate system as shown in following equation.

$$ED = \sqrt{\sum_{i=1}^{k} (x_i - y_i)^2}$$

where ED is Euclidean distance. In a two dimensional, the distance is just between two points (xi, yi) in xy plane space where (i D 1; 2 : : : ; k) be data points [27]. In similar way Manhattan formula and Minkowski formula is also used for calculating the distance. We will train and tested k nearest neighbor algorithm on Gaussian distribution for pattern recognition as given in research paper.

Step 4: Training and Testing Models

Once all models get build then will train each model for given dataset by passing dependent(X) and independent matrices(Y) as parameters to respective models.

Here X matrix is referred as independent variables training matrix which contains around 800 records.

And Y matrix is dependent variable's training matrix which contains around 800 records.

Once model is train then we will test it for another X and Y matrix of independent and dependent variables which contains around 400 record's each.

Step 5: Design Web-App And Integration With ML Models

Once models are ready then we will start designing website whose main task will be taking online test to predict results for Prakriti analysis.

For this we will use following techniques :

1. **Front End:**

   For UI design we used normal HTML to design components of site while for decoration purpose we used CSS and Bootstrap and to perform different actions we will use JavaScript.

2. **Backend:**

   To integrate ML models with web app we used flask tool which works on python language.

❑ Chapter 5: Implementation Results and Performance Analysis

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.

For better performance of models, we will try different methods to increase the accuracy of models like Bagging, boosting and hyperparameter tuning etc. Which helps to us to select best parameter values during implementation of models.

Till now we have just implemented 2 models

Performance analysis of Naïve Bayes and SVM Model :

We can generate classification report for our model using inbuild python libraries as shown in following screenshots which gives information about different values like Precision, Recall, F-Score, etc.

1. PRECISION (POSITIVE PREDICTED VALUE)

Precision is used to limit the number of false positives. It checks that how often a classifier predicts the positive results. It is calculated as number of correct positive predictions divided by total number of positive predictions

can be calculated as,

Precision = TP/TP + FP

2. RECALL (TRUE POSITIVE RATE)

It measures the sensitivity and is calculated as number of correct positive predictions divided by total number of positives.

Recall = TP/TP + FN

3. F-SCORE

F-Score is a measure of a test accuracy and is defined as the weighted harmonic mean of the precision and recall of the test.

F - Score = 2 * (Precision * Recall)/(Precision + Recall)

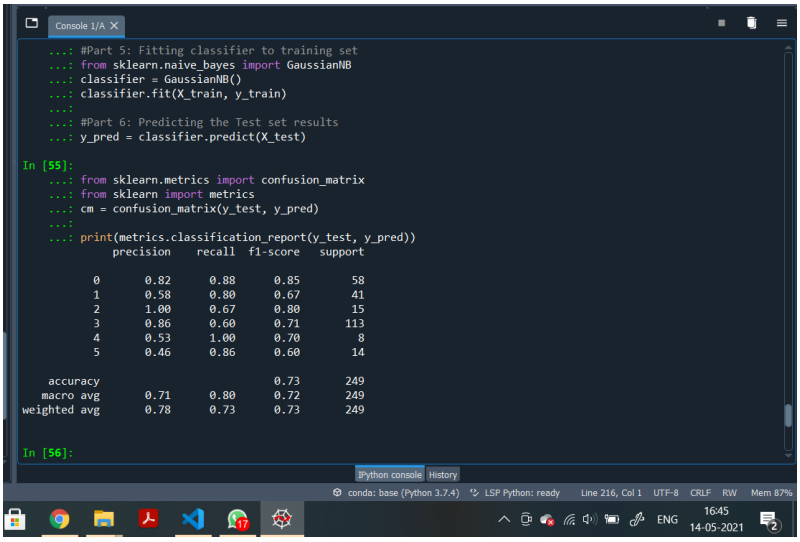
4. Accuracy

Accuracy is a measure used for machine learning models to determine winning model for identifying relationships and patterns between variables in a dataset based supervised data. Accuracy can be calculated as,

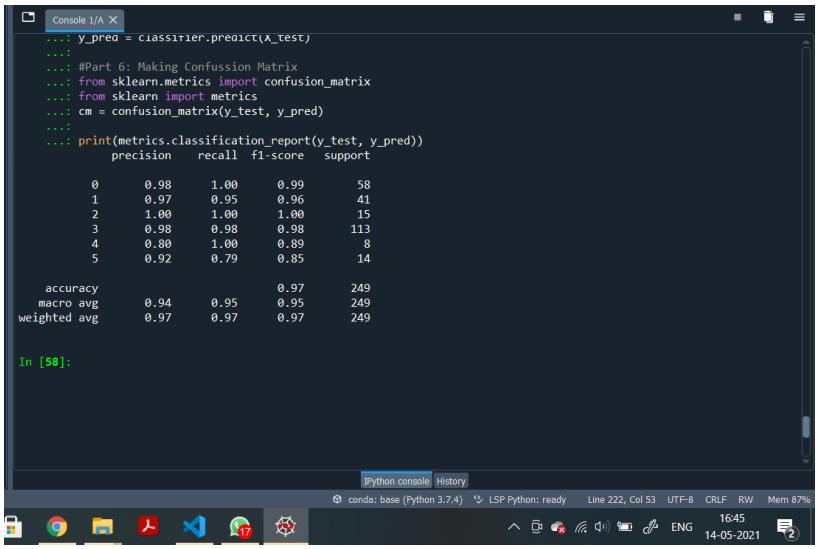
Accuracy = TP + TN / TP + FP + TN + FN

Classification Report :

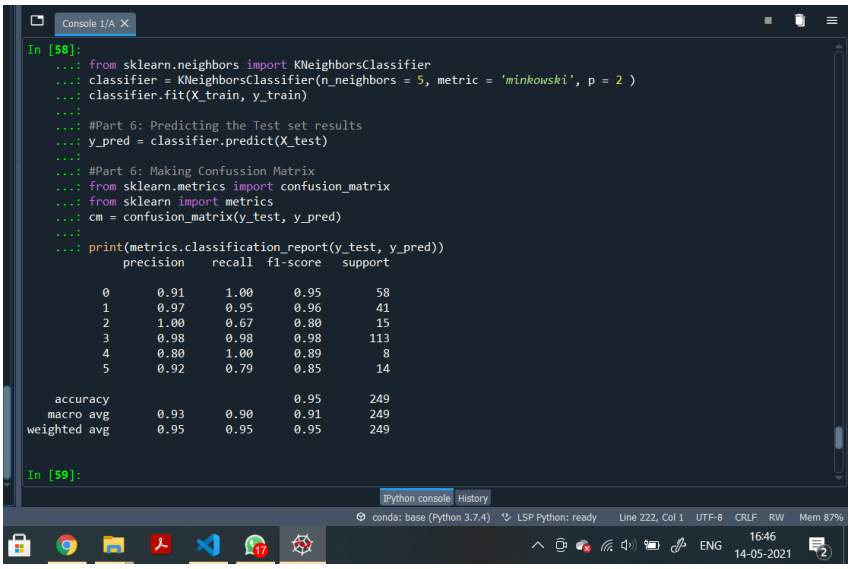
1. Naïve Bayes Model -



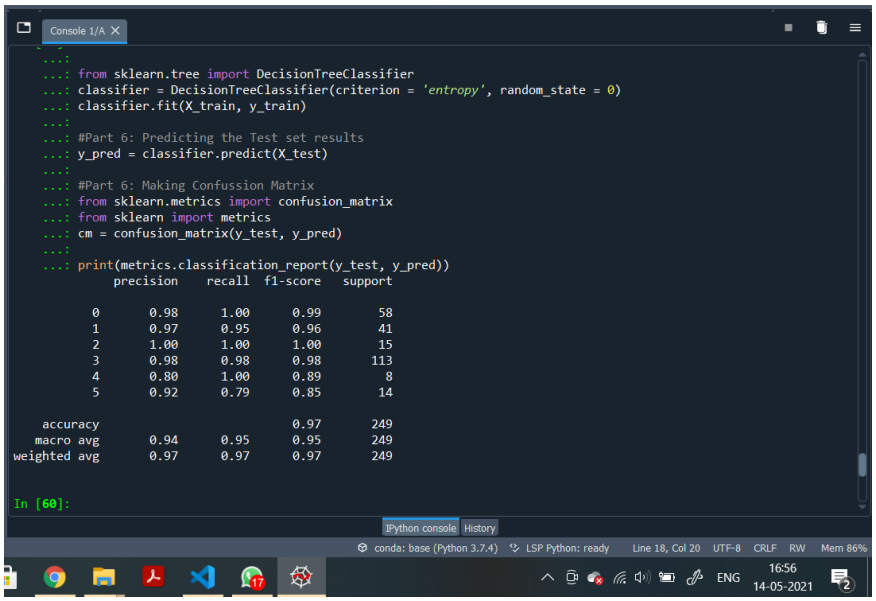
2. SVM Model :



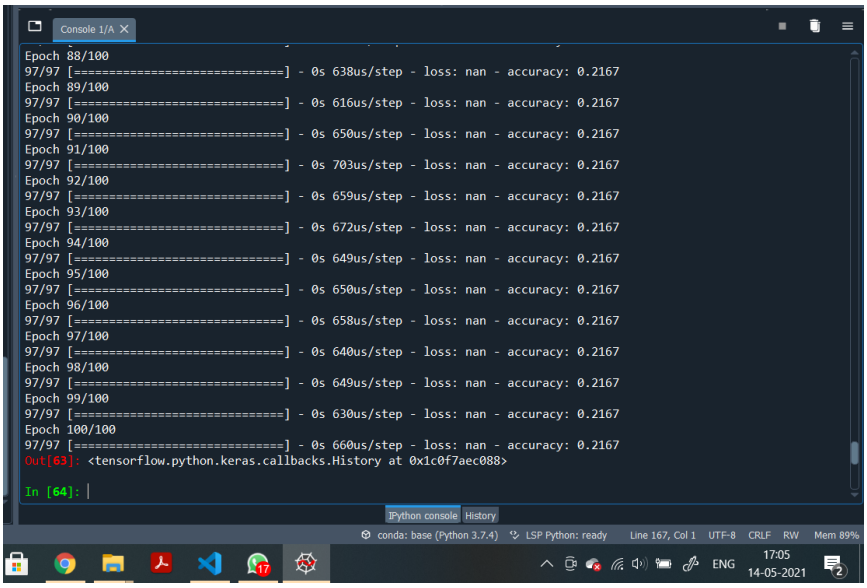
3. KNN Model :



4. Decision Tree :



5. ANN :



Confusion Matrix :

1. **Naïve Bayes Model -**

cm - NumPy object array

	0	1	2	3	4	5
0	51	0	0	5	0	2
1	1	33	0	6	0	1
2	0	0	10	0	5	0
3	10	23	0	68	1	11
4	0	0	0	0	8	0
5	0	1	0	0	1	12

2. SVM -

cm - NumPy object array

	0	1	2	3	4	5
0	58	0	0	0	0	0
1	1	39	0	1	0	0
2	0	0	15	0	0	0
3	0	0	0	111	1	1
4	0	0	0	0	8	0
5	0	1	0	1	1	11

3. KNN -

cm - NumPy object array

	0	1	2	3	4	5
0	58	0	0	0	0	0
1	1	39	0	1	0	0
2	5	0	10	0	0	0
3	0	0	0	111	1	1
4	0	0	0	0	8	0
5	0	1	0	1	1	11

4. Decision Tree -

cm - NumPy object array

	0	1	2	3	4	5
0	58	0	0	0	0	0
1	1	39	0	1	0	0
2	0	0	15	0	0	0
3	0	0	0	111	1	1
4	0	0	0	0	8	0
5	0	1	0	1	1	11

Here,

0 Indicates VATT(V)

1 Indicates PITT(P)

2 Indicates KAFF(K)

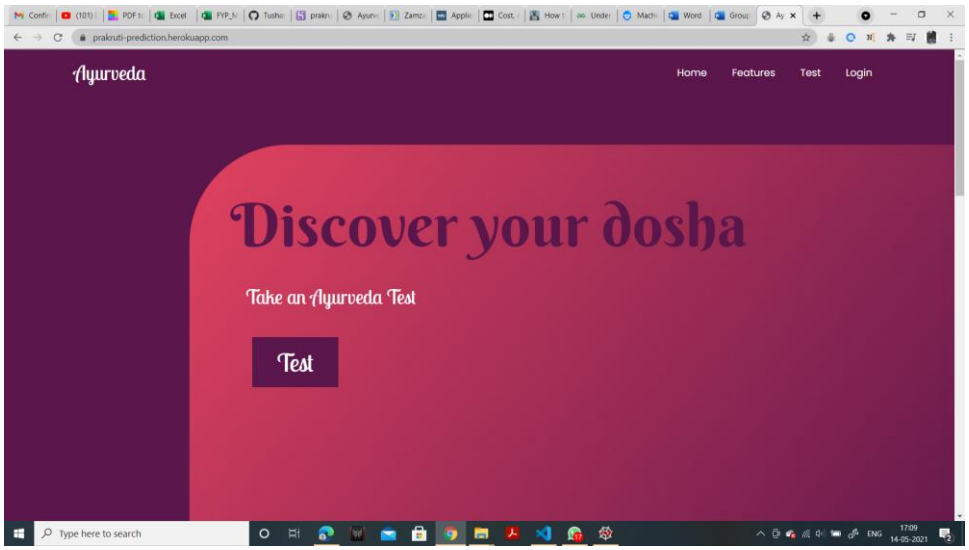
3 Indicates VATT-PITT(V-P)

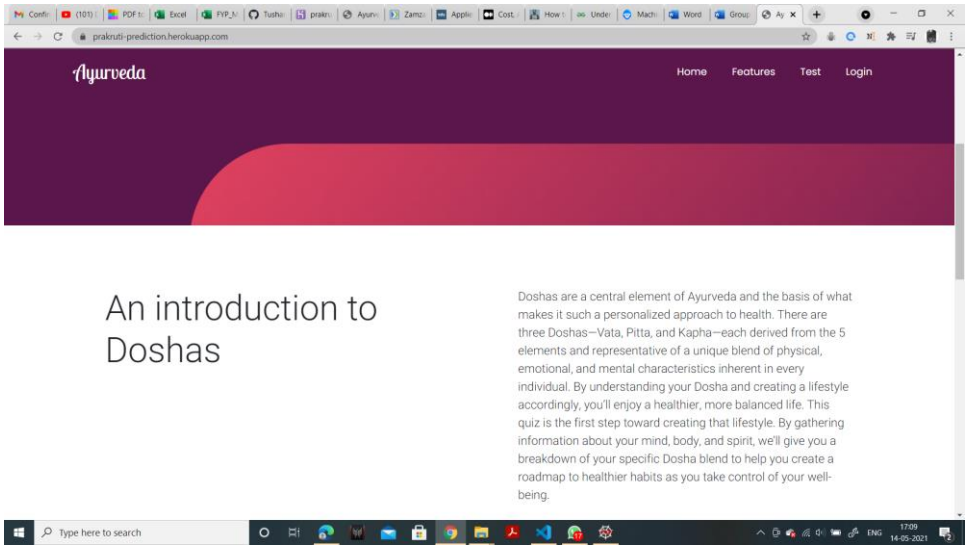
4 Indicates PITT-KAFF(P-K)

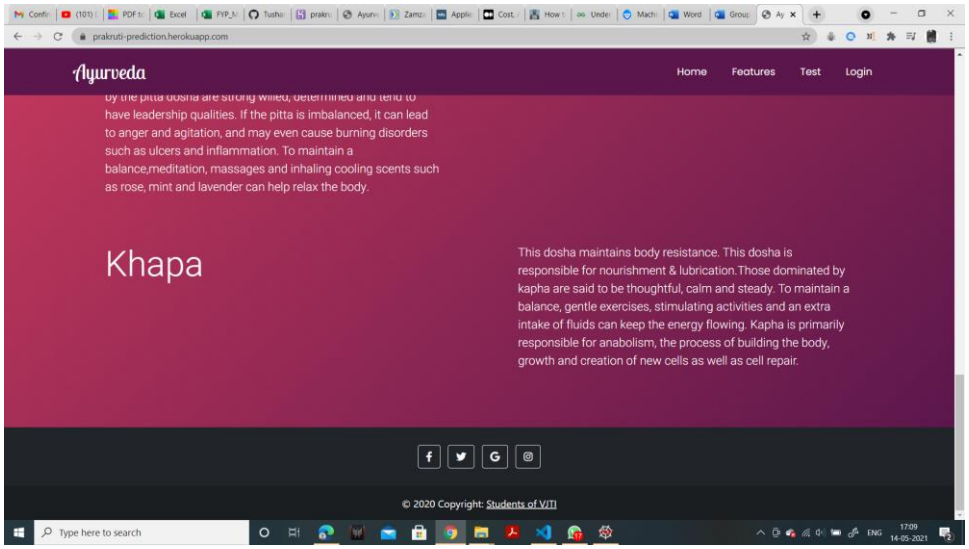
5 Indicates VATT-KAFF(V-K)

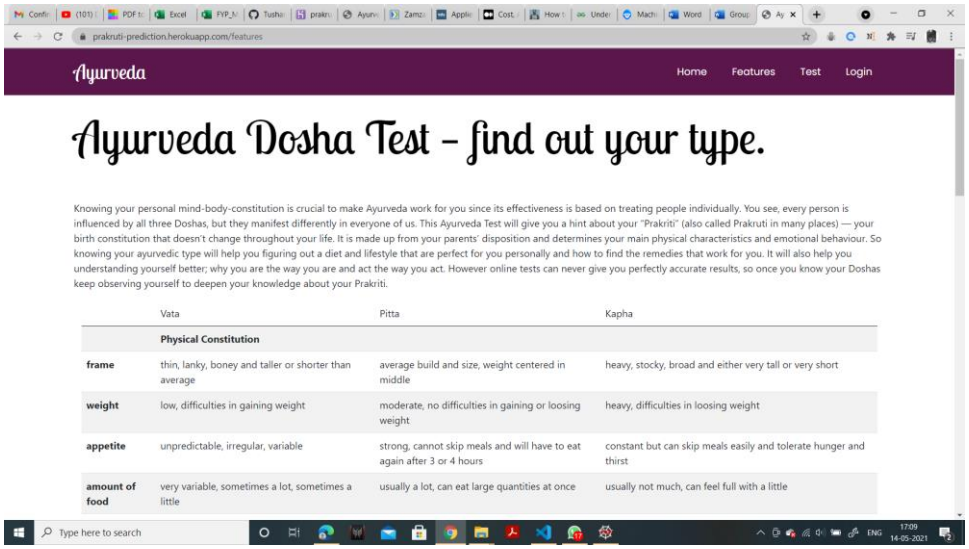
From above screenshots we can say that SVM and Decision model is really performing well compare to other Models.

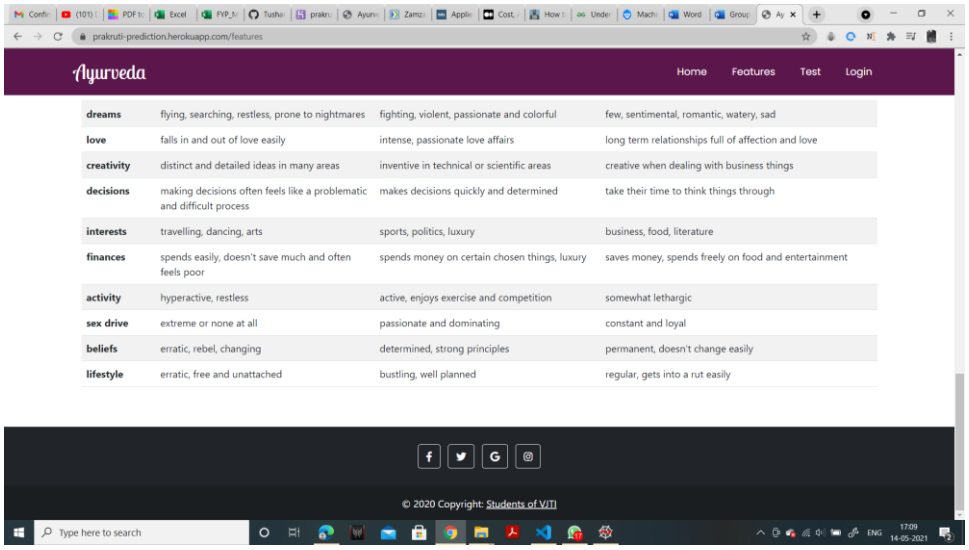
❑ Chapter 6: UI Design

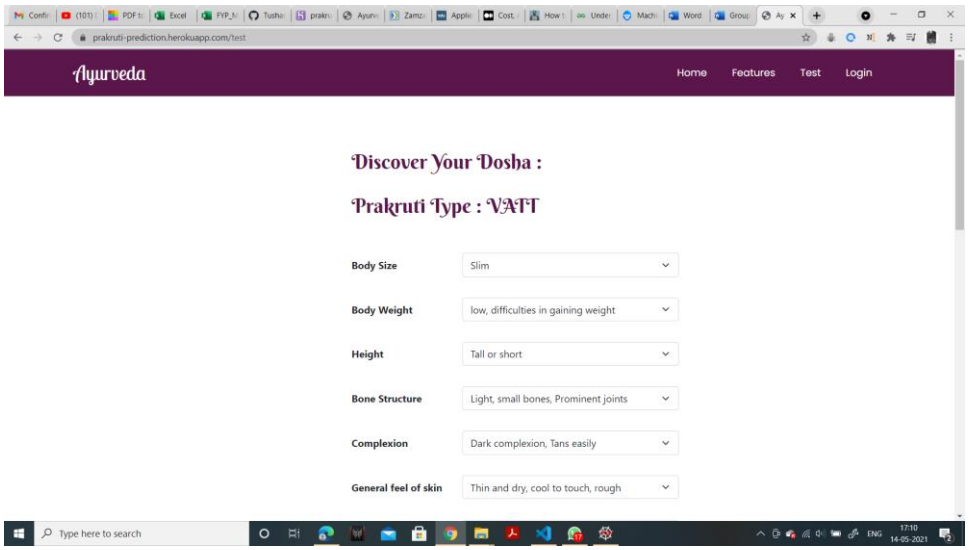


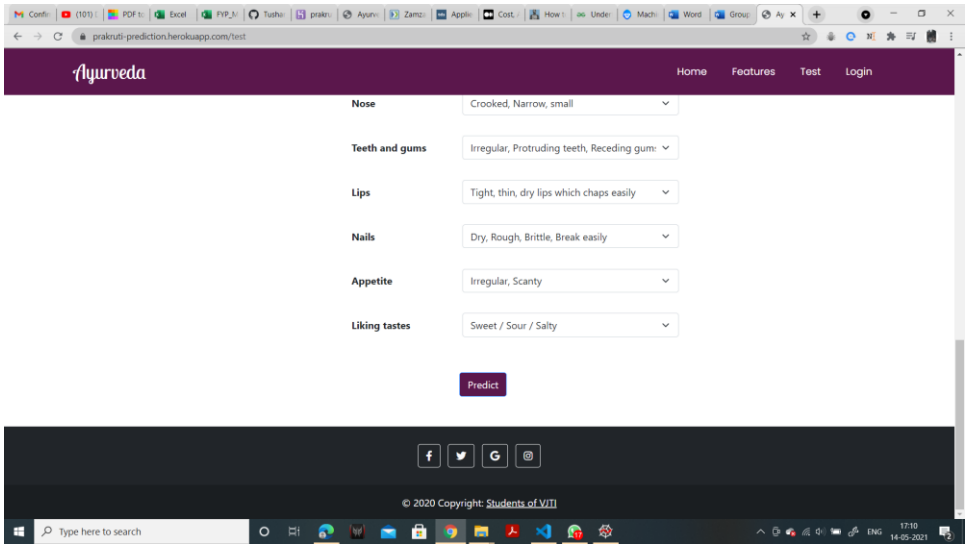


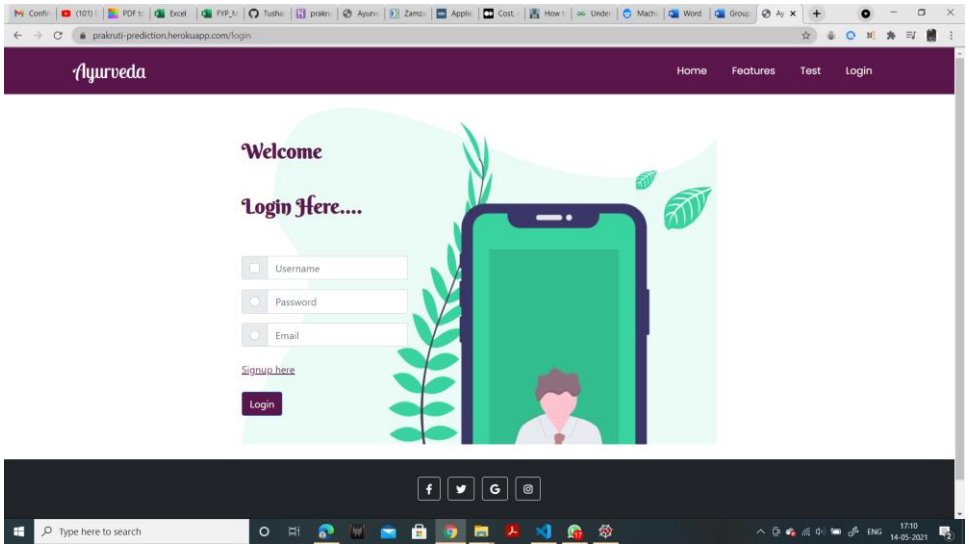


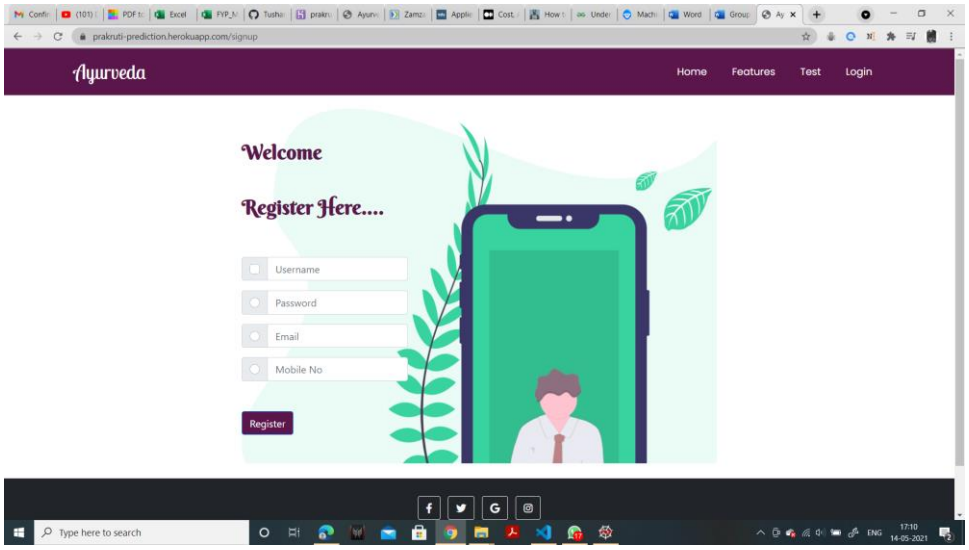

prakruti-prediction.herokuapp.com/features
Ayurveda
Home
Features
Test
Login
Ayurveda Dosha Test – find out your type.
Knowing your personal mind-body-constitution is crucial to make Ayurveda work for you since its effectiveness is based on treating people individually. You see, every person is influenced by all three Doshas, but they manifest differently in everyone of us. This Ayurveda Test will give you a hint about your "Prakriti" (also called Prakruti in many places) — your birth constitution that doesn't change throughout your life. It is made up from your parents' disposition and determines your main physical characteristics and emotional behaviour. So knowing your ayurvedic type will help you figuring out a diet and lifestyle that are perfect for you personally and how to find the remedies that work for you. It will also help you understanding yourself better; why you are the way you are and act the way you act. However online tests can never give you perfectly accurate results, so once you know your Doshas keep observing yourself to deepen your knowledge about your Prakriti.
Vata
Pitta
Kapha
Physical Constitution
frame
thin, lanky, boney and taller or shorter than average
average build and size, weight centered in middle
heavy, stocky, broad and either very tall or very short
weight
low, difficulties in gaining weight
moderate, no difficulties in gaining or loosing weight
heavy, difficulties in loosing weight
appetite
unpredictable, irregular, variable
strong, cannot skip meals and will have to eat again after 3 or 4 hours
constant but can skip meals easily and tolerate hunger and thirst
amount of food
very variable, sometimes a lot, sometimes a little
usually a lot, can eat large quantities at once
usually not much, can feel full with a little
Type here to search
17:09
14-05-2021


prakruti-prediction.herokuapp.com/features
Ayurveda
Home
Features
Test
Login
dreams
flying, searching, restless, prone to nightmares
fighting, violent, passionate and colorful
few, sentimental, romantic, watery, sad
love
falls in and out of love easily
intense, passionate love affairs
long term relationships full of affection and love
creativity
distinct and detailed ideas in many areas
inventive in technical or scientific areas
creative when dealing with business things
decisions
making decisions often feels like a problematic and difficult process
makes decisions quickly and determined
take their time to think things through
interests
travelling, dancing, arts
sports, politics, luxury
business, food, literature
finances
spends easily, doesn't save much and often feels poor
spends money on certain chosen things, luxury
saves money, spends freely on food and entertainment
activity
hyperactive, restless
active, enjoys exercise and competition
somewhat lethargic
sex drive
extreme or none at all
passionate and dominating
constant and loyal
beliefs
erratic, rebel, changing
determined, strong principles
permanent, doesn't change easily
lifestyle
erratic, free and unattached
bustling, well planned
regular, gets into a rut easily
© 2020 Copyright: Students of VIT
Type here to search
17:09
14-05-2021


prakruti-prediction.herokuapp.com/test
Ayurveda
Home
Features
Test
Login
Discover Your Dosha :
Prakruti Type : VATT
Body Size
Slim
Body Weight
low, difficulties in gaining weight
Height
Tall or short
Bone Structure
Light, small bones, Prominent joints
Complexion
Dark complexion, Tans easily
General feel of skin
Thin and dry, cool to touch, rough
Type here to search
17:10
14-05-2021


prakruti-prediction.herokuapp.com/test
Ayurveda
Home
Features
Test
Login
Nose
Crooked, Narrow, small
Teeth and gums
Irregular, Protruding teeth, Receding gum:
Lips
Tight, thin, dry lips which chaps easily
Nails
Dry, Rough, Brittle, Break easily
Appetite
Irregular, Scanty
Liking tastes
Sweet / Sour / Salty
Predict
© 2020 Copyright: Students of VJTI
Type here to search
17:10
14-05-2021


prakruti-prediction.herokuapp.com/login
Ayurveda
Home
Features
Test
Login
Welcome
Login Here....
Username
Password
Email
Signup here
Login
Type here to search
17:10
14-05-2021


prakruti-prediction.herokuapp.com/signup
Ayurveda
Home
Features
Test
Login
Welcome
Register Here....
Username
Password
Email
Mobile No
Register
Type here to search
17:10
14-05-2021

❑ Chapter 7: Conclusion

The main goal of this project is to predict Prakriti of human beings based human body constituents using different machine learning models and analyze their performance on various parameters. To achieve this, we First design our own dataset by taking survey on www.surveycorcle.com. After that, we analyze prakriti of each record with the help of Ayurveda specialist and finally now we are building different machine leaning models K-nearest neighbor, artificial neural networks, support vector machine, Naive Bayes, decision tree without hyper parameter tuning. Till now we have implemented half machine learning models which we have already integrated with our web app and in future we are planning to deploy our web site on Heroku platform where peoples can take their test to see their prakriti in less than a minute.

Reference's :

1. IEEE Research paper on Predicting Ayurveda-Based Constituent Balancing in Human Body Using Machine Learning Methods - https://ieeexplore.ieee.org/document/9057416
2. https://www.kaggle.com/learn/intro-to-machine-learning
3. https://www.planetayurveda.com/prakriti-analysis
4. https://en.wikipedia.org/wiki/Ayurveda
5. https://www.w3schools.com/
6. https://dashboard.heroku.com/apps