ENHANCING AYURVEDIC DIAGNOSTICS THROUGH MACHINE LEARNING VIA INTEGRTION OF SMART HEALTH MONITORING DATA

THESIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE AWARD OF DEGREE OF

MASTER OF TECHNOLOGY

in

ARTIFICIAL INTELLIENCE

by

YATI PIPLANI

2K22/AFI/29

under the supervision of

Dr. PAWAN SINGH MEHRA

(Assistant Professor)

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-110042, India

May 2024

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

CANDIDATE'S DECLARATION

I, **Yati Piplani**, Roll No. 2K22/AFI/29 student of M. Tech (Artificial Intelligence), hereby declare that the Project Dissertation titled **"Enhancing Ayurvedic Diagnostics through Machine Learning via Integration of Smart Health Monitoring Data"** which is being submitted by me to Delhi Technological University, Delhi, in partial fulfillment of requirements for the degree of Master of Technology in Artificial Intelligence is a legitimate record of my work and is not copied from any source. The work contained in this report has not been submitted at any other University/Institution for the award of any degree.

Place: Delhi Yati Piplani

Date:13 May 2024 (2K22/AFI/29)

This is to certify that the student has incorporated all the corrections suggested by the examiners in the thesis and the statement made by the candidate is correct to the best of my knowledge.

Signature of Supervisor **Signature of External Examiner**

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

CERTIFICATE

I, hereby certify that the Project titled **"Enhancing Ayurvedic Diagnostics through Machine Learning via Integration of Smart Health Monitoring Data,"** submitted by Yati Piplani, Roll No. 2K22/AFI/29, Department of Computer Science & Engineering, Delhi Technological University, Delhi in partial fulfillment of the requirement for the award of degree of M. Tech in Artificial Intelligence is a genuine record of the project work carried out by the student under my supervision. To the best of my knowledge this work has not been submitted in part or full for any Degree to this University or elsewhere.

Place: Delhi — Dr. Pawan Singh Mehra

Date: 13 May 2024 — (Assistant Professor)

ABSTRACT

Incorporating machine learning (ML) and smart health monitoring with Ayurvedic diagnostics is a novel approach to improving our traditional medical practices. To level up the recognition and categorization of the three Ayurvedic doshas—Pitta, Kapha, and Vata—the study uses an unsupervised learning approach to train health-related data from wearable devices like smartwatches and smartphones. With the help of useful data gathered from smart health devices—such as heart rate, physical activity, and sleep patterns—this study aims to improve the quality of Ayurvedic assessment and personalize treatment plans. This study took sleep data from the collected one to analyze the Prakriti.

Using K-means clustering on sensored collected data, we are able to find patterns associated with it. This method makes use of unsupervised learning's capabilities to recognize undetectable patterns in datasets without predetermined labeling. The main dataset that we have utilized is the information gathered from smart devices that people usually wear and use to continuously monitor different physiological characteristics. After doing thorough preprocessing to guarantee the consistency and validity of our approach, EDA is performed to identify major attributes of Ayurvedic diagnosis.

Based on collected information, the study's findings suggest that machine learning can distinguish between the three human constitutions with good accuracy; certain patterns are also categorized by clusters that are interrelated to the Ayurveda understanding of doshas. We have identified that by using machine learning methods, it is possible to identify the distinct patterns exhibited by each dosha. For example, those with a higher Vata characteristic do not sleep for longer periods and have more peaceful sleep patterns, and then some people with a higher Kapha characteristic.

We offer a novel framework that combines Ayurveda and technology to produce a more comprehensive, data-driven approach to diagnosing health issues and providing quick resolution with the touch of tradition. This research has advanced the field. It draws attention to how ML may improve the precision and customization of Ayurvedic treatments, making them more usable and relevant in the current digital time. Also, in order to globalize subsequent research could also take an effort that points more on increasing the datasets with more demographic scope, with more number of parameters, with more health metrics, and also standardized dataset for enhancing the generalizability in this expanding field. Also, we can use other machine learning approaches that could further enhance this diagnostic process.

In the conclusion it has been seen by extensive literature review and experiments that there is potential to reform the traditional healthcare field with the combination of conventional procedures, artificial intelligence, and smart health monitoring devices. As per Ayurvedic principles, this combination could result in improved health outcomes and a deeper understanding of individual health characteristics by offering a more precise, personalized, and preventive approach to health management.

ACKNOWLEDGMENT

I am extremely grateful to my project guide, **Dr. Pawan Singh Mehra,** Assistant Professor, Department of Computer Science and Engineering, Delhi Technological University, Delhi for providing valuable guidance and being a constant source of inspiration throughout my research. I will always be indebted to him for the extensive support and encouragement he provided.

I am highly indebted to the panel faculties during all the progress evaluations for their guidance, constant supervision and for motivating me to complete my work. They helped me throughout by giving new ideas, providing necessary information, and pushing me forward to complete the work.

Yati Piplani
(2K22/AFI/29)

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVATION

AI	Artificial Intelligence
ML	Machine Learning
DL	Deep Learning
LR	Logistic Regression
SVM	Support Vector Machine
DT	Decision Tree
KNN	K Nearest Neighbor
NB	Naïve Bayes
DT	Decision Tree
RF	Random Forest
DBN	Deep Belief Network
FFNN	Feed Forward Neural Network
ANN	Artificial Neural Network
MLP	Multilayer Perceptron
SGD	Stochastic Gradient Descent
DNN	Deep Neural Network
IOT	Internet of Things
VP	Vata Pitta
PV	Pitta Vata
VK	Vata Kapha
PVK	Pitta Vata Kapha
TN	True Negative
TP	True Positive
FN	False Negative

FP	False Positive
AUC	Area under Curve
ROC	Receiver Operating Characteristic Curve
ASGF	Ayurvedic Smart Generalized Framework

CHAPTER 1

INTRODUCTION

1.1 Overview of Ayurveda

Ayurveda [1] is an ancient medical system which is originated from India and has been a root of our holistic health culture by thousands of years. This traditional system emphasizes mainly upon the balance of body, mind, and spirit which helps to prevent and treat illness. It is not like the western medicine which is always focused on treating symptoms not the cause. But on the other hand ayurveda always helped to eliminate the root cause of disease through personalized care and with the help of panchtatva principles. The practices followed in ayurveda includes pulse diagnosis, where practitioners measure the pulse rate to captures the body malfunctions. Another method employed is tongue diagnosis which involves examining the tongue for signs of various health issues according to their color. This holistic approach allows our doctors to provide a deeply personalized medical practice which aim is not just to cure illness but by also maintaining overall health and well-being. But fundamental concept of the ayurveda lies within the three doshas [2]: Vata, Pitta, and Kapha. These doshas are basic energies which are found throughout the human body and mind an helps in governing all physical and mental processes and also in providing every living being with an individual care for healthy living lifestyle.

FIGURE 1.1: Human Body Constituents

The fig. 1.1 shows the three human body constituents and is defined below:

- **Vata** is a combination of its corresponding qualities which are basically reflecting the elements of space and air. It helps to control the movement and is considered as a main force behind the nervous system's working, breathing, and circulation. Individuals which are dominant in Vata Doshas are quick-thinkers, thin, and fast-moving.

- **Pitta** represents the elements of fire and water. It governs digestion and metabolism in the organ and tissue systems. Those with a Pitta constitution are usually seen as competitive, fiery, and capable of strong intellectual digestion.

- **Kapha** combines the elements of earth and water. It is responsible for strength, stability, and moisture. Kapha balances the body's structure and fluidity and is characterized by attributes of calmness, solidity, and endurance.

An Ayurvedic doctor uses these different constituents to assess imbalance by checking the disbalance and then prescribe treatments accordingly which ranges from changes in diet to herbal therapies. With the understanding of an individual's dominant dosha and imbalances present ayurvedic medicine aimed to restore balance and harmony inside our body and lead us to spend more healthier and cheerful life. By this good understanding of one's constitution used the personalized approach that defines Ayurvedic medicine.
In Ayurvedic practice, individuals are categorized based on the predominance and balance of their dominant doshas that helps in understanding their natural constitution and thereby providing specific health recommendations. Table 1.1 briefly describes the Doshic constitutions based on the dominance of the Doshas and is explained below:

- **Single Dosha Dominance:** Some individuals primarily contain characteristics of one dosha only which are either Vata (V), Pitta (P), or Kapha (K). This single dominance develops their overall health tendencies and personality traits.

- **Dual Dosha Dominance:** Some persons may have two relatively balanced doshas but one of them still predominates the other. These combinations can be Vata-Pitta (VP), Pitta-Vata (PV), or Vata-Kapha (VK) and each one of them are picturing a unique mixture of characteristics from the available doshas.

- **Tri-Dosha Balance:** It is a rarer constitution in which all the three doshas i.e. Vata, Pitta, and Kapha are in almost equal proportion and it is denoted by VPK. This balance is considered to be ideal and it also leads to a balanced and healthy lifestyle with a very rare health problem.

TABLE 1.1: Doshic Constitutions based on Dominant Doshas

Type	Description
V, P, K	Predominant in Single Body Constitution (Vata, Pitta, or Kapha)

VP, VK, PV	Two Body Constitution in relatively equal proportions, in which one is predominating
VPK	Body Constitution in almost half and half proportion

With the help of Table 1.2, we have provided a detailed overview of different characteristics associated with each dosha and also by covering the key aspects like physical attributes like body size and weight to psychological behaviors such as emotional and mental characteristics. This classification not only helps in the diagnosing the health imbalances but also in tailoring personalized lifestyle and treatment plans that align with an individual's body constitution. With the understanding of these doshic characteristics, is important for practitioners and patients both because it forms the foundation to achieve balance and well-being through Ayurvedic principles.

TABLE 1.2: Characteristics associated with each Doshas

Observation	Vata	Pitta	Kapha
Body size	Thin	Average	Broad
Body weight	Light	Normal	Heavy
Cheeks	Creased/sunken	Even and smooth	Full, soft
Face shape/chin	Narrow, sharp	Wedge-shaped	Broad with a double chin
Eyes	Small, restless, dark	Penetrating, light-sensitive	Large, gentle, blue
Nose	Asymmetrical, irregular	Sharp, reddish tip	Rounded, stubby
Lips	Chapped, dark	Bright red, sensitive	Soft, pale
Teeth	Protruding, sparse	Average, sensitive gums	Strong, white
Skin	Rough, cool, pale	Soft, warm, pink	Smooth, cool, pale
Hair	Sparse, tangled, brittle	Fine, prone to balding	Lush, wavy
Appetite	Variable, low	Intense, immediate	Consistent, moderate
Digestion	Erratic, gassy	Fast, overheating	Slow, mucous-heavy
Thirst	Variable	Excessive	Limited
Emotions	Anxious, uncertain	Irritable, passionate	Peaceful, possessive
Mind	Agitated	Decisive	Steady
Intellect	Quick, impractical	Precise	Methodical, thorough
Speech	Fast, indistinct	Articulate, intense	Measured, dull
Voice	Faint, raspy	Vibrant	Rich, resonant

1.2 Machine Learning in Healthcare

In the rising era the machine learning (ML) [3] has transformed the healthcare system, particularly in the field of disease diagnosis. ML uses various techniques from Supervised to reinforcement learning that are helpful in utilizing large datasets for analysis purposes which is not handled by standard statistical approaches. It helps to analyze new complex patterns and correlations that human practitioners may miss during observation. ML allows us to see a big view of applications using various datasets available and constructed. This feature is critical for early disease detection, diagnosis, and potentially forecasting patient outcomes.
Machine learning models process and also interpret a big range of data types like text, images, videos, etc in their training periods also including radiology and genetic information as well as electronic health records. For example, ML algorithms can evaluate thousands of radiographic pictures and discover anomalies such as tumors, fractures, or infections like pneumonia with high accuracy and precision. The strength of ML lies in its ability to continuously learn and improve from more data over time.

Specific machine learning techniques that hold relevance in healthcare include:

- **Classification Algorithms:** These [4] are used to categorize data into predefined classes. For ex: In the context of healthcare it is helpful in determining whether a particular image of x-ray has fractured or not.

- **Regression Models:** It is useful for predicting continuous number and also regression models [5] can forecast things like the days of hospital stays or patients' response levels to different treatments based on historical data present.

- **Neural Networks:** It is a early deep learning model which is influenced by how to brain neuron works. In the healthcare field a subset of neural networks [6], has been influential in advancing areas such as medical image analysis. These models can learn to perform tasks such as segmenting tissues, identifying disease markers and even suggest possible diagnoses.

- **Clustering Algorithms:** These algorithms are used to group datapoints in such a way that data's with the similar nature are in the same group (or cluster) [7]. In healthcare, clustering can identify subgroups within patient populations based on similarities in their medical records, which can be useful in personalized medicine strategies and understanding patient subtypes.

1.3 Integration of Machine Learning with Ayurvedic Diagnostics

The application of machine learning (ML) in Ayurvedic diagnostics is a promising area in advancing the traditional medicinal field. It is seen that with different experiments and research that took place by combining machine learning with smart health monitoring data can give Ayurvedic practitioners a great way to improve diagnosis process. Smart health devices that monitor activity levels, and other features like sleep, calorie and heart is a great sources of useful data and when it is evaluated using machine learning algorithms and also it provides a more accurate and precise picture of an individual’s health.

- **Enhanced Diagnostic Accuracy:** Machine learning can help to train complex and varied data sets collected from smart health devices to identify patterns and trends that may not be possible or difficult with traditional practices alone. For example, ML algorithms can detect a small changes in heart rate variability or activity patterns that correlate with imbalances in the doshas (Vata, Pitta, Kapha) very accurately and fast. This capability could lead to earlier and more accurate identification of doshic imbalances, enhancing the precision of Ayurvedic diagnostics.

- **Personalized Treatment Adjustments:** The integration of ML allows the continuous analysis of health metrics of patients that helps Ayurvedic practitioners with insights that can prompt timely adjustments to treatment plans. As ML models learn from new data therefore it can also predict how a patient might respond to certain treatments which enables practitioners to personalize healthcare methods based on real-time data. This adaptability help to improve the treatment outcomes and patient satisfaction by ensuring that interventions are finely tuned to changing health status of persons.

- **Scalability of Ayurvedic Practices:** It is seen that Machine learning faces one of the major challenges in traditional medicine which is scalability issue. By taking care of this part of the diagnostic process, ML enables practitioners to handle larger volumes of patients without compromising the quality of care. Moreover, ML algorithms can help in the standardization of diagnostic criteria and treatment protocols in Ayurveda which in return makes it easier to train new practitioners and expand the reach of Ayurvedic medicine to reach globally.

1.4 Rise of Smart Health Monitoring

The increase in the smart health monitoring systems [8] is making a significant evolution in the way of collection and usage of health data. These smart systems provides a range of devices which are mostly wearables and are also very classy and user-friendly. The common wearable devices which includes fitness trackers, smartwatches. Also they consist of unique and well equipped garments that are modified with sensors and can continuously gather health-related data. These devices mostly track heart rate, sleep patterns, physical activity. Also, sometimes they are able to track more advanced parameters like blood oxygen levels and also skin activity like amount of stress, heat or even sweating.

The data that are collected with the help of these devices offers a variety of information that can be used to analyzed and assess human health and goodness in real-time. For example, heart rate monitors can help track cardiovascular health and also it can detect defects that may also indicate stress or cardiac issues. Moreover, Sleep trackers present inside the device provides the insights into sleep quality, duration, and patterns which are essential for understanding the impact of sleep on overall health. Similarly, activity trackers monitor the amount and intensity of physical activity which is very much crucial for managing weight and enhancing physical fitness.

The potential of smart wearables lies in its vast amount of health metrics that they are collecting from individual human beings . These data can add a cherry on top of the traditional diagnostic methods by providing steady and accurate measurements that are not only limited to the questionaries dataset but also to the other factors as well. For example while doing continuous monitoring it can reveal trends and patterns that a normal clinical tests might miss which includes the important information such as heart rate variability or the sleep stage distributions in real time.

Additionally, the integration of the artificial intelligence technology with traditional healthcare methods provide us a more well versed approaches for the diagnosis and treatment. In clinics the physicians can use data analyzed from wearables with the help of ML models make them to take more informed decisions and also monitor the effectiveness of treatments over time to time. Moreover, they can also adjust prescriptions that is based on objective data inputs which are provided to them. For traditional practices like Ayurveda these integration with smart health data enhance the accuracy of dosha assessments and refine personalized treatment plans.

In summary we conclude that smart health monitoring technologies are helping out in reshaping the healthcare system by bridging the gap between traditional diagnostic practices with advance and data-driven approaches. This integration can not only enhance patient care but also empower individuals to take an active participation in managing their health by promising a future where personalized and preventive healthcare becomes a part of day to day lifestyle.

1.5 Primary Hypothesis

1. **Existing solutions are mostly revolving around supervised learning methods:** In the current situation of healthcare diagnostics enhanced by machine learning, supervised learning methods [9] predominate so much. These methods require labeled datasets, where the input data (such as images, patient symptoms, or lab results) are paired with predefined

outputs (like diagnostic categories or disease markers). This reliance on supervised learning is particularly evident in areas requiring high accuracy, such as disease detection from medical images or prediction of specific health outcomes based on clinical data.

However, this approach has limitations, particularly in fields like Ayurveda, where the diagnostic categories are not always clearly defined or universally agreed upon, and where symptoms and their interpretations can be highly subjective. Supervised learning models are only as good as the data they are trained on, which means they might not capture the holistic essence of Ayurvedic diagnostics that consider physical, emotional, and spiritual health collectively.

2. **Existing solutions are mostly dependent on questionnaires' datasets:** Many existing diagnostic tools in both conventional and English medicine rely heavily on data collected through questionnaires. These tools use responses from structured questionnaires to identify symptoms, lifestyle factors, and patient histories. We know that this method provides structured data that is easy to analyze but it can be inherently biased and caries limitations of these reported data. For example, the accuracy of the responses can be influenced by a patient's memory, their understanding of the questions, or even their willingness to share personal health details. In Ayurveda, reliance on questionnaire data might not fully capture the overall observations that an practitioner can make during a physical examination or through direct interaction. Furthermore, questionnaires typically do not capture real-time physiological data, which can be crucial for understanding a patient's current state of health.

1.6 Objectives

1. **To conduct a literature review on improvising the Ayurveda Doshas identification and classification using various machine learning techniques:** This objective involves a comprehensive analysis of existing research and methodologies that apply machine learning to the diagnosis and classification of Ayurvedic Doshas. The focus should be on evaluating how different machine learning approaches, particularly supervised and semi-supervised techniques, have been utilized and their effectiveness in capturing the complexities of Dosha imbalances. This review will also identify gaps in current methodologies, such as over-reliance on supervised learning and limited use of real-time physiological data, setting the stage for the introduction of innovative approaches, like unsupervised learning, in this traditional healthcare system.

2. **To develop a novel solution towards the identification of these Doshas using an unsupervised learning method, which is not much explored in this area:** Given the limitations identified in traditional and current machine learning applications to Ayurvedic diagnostics, this objective seeks to pioneer the use of unsupervised learning [10] techniques.

Unsupervised learning does not require labeled datasets; instead, it identifies patterns and relationships in the data on its own. This approach is particularly suitable for Ayurvedic diagnostics, where explicit labels (like specific disease identifiers in Western medicine) might not always be available or clearly defined. Developing a model that can cluster and interpret complex health data without predefined categories could revolutionize how practitioners understand and treat Dosha imbalances.

3. **To make use of the right datasets and to perform accurate data analysis methods to ensure proper outcomes:** This objective emphasizes the critical importance of data quality and analysis techniques in achieving reliable diagnostic results. It involves selecting the most relevant and comprehensive datasets, potentially integrating traditional questionnaire data with real-time data collected from smart health monitoring devices. The analysis methods chosen must be capable of handling the heterogeneity and volume of the data, ensuring that the insights generated are both accurate and applicable in a real-world clinical setting. Additionally, this objective includes validating the machine learning models developed to ensure they are robust, scalable, and capable of providing actionable diagnostic information.

1.7 Research Questions

Q1. How can Ayurvedic interpretations be integrated with data collected from wearable devices?

Q2. How effectively can unsupervised machine learning techniques identify different Prakriti types from health data?

Q3. How can this method be utilized in individualized health management within the Ayurvedic practice?

CHAPTER 2

RELATED WORK

2.1 Literature Overview

This section will cover into novel studies that are undertaken in recent years by also providing essential insights into the present scenarios of tradition and technological advancements. In the past few years, Artificial Intelligence [11] has become instrumental in many real-life problems like healthcare, social networks, recommendation systems, fake news and negative stances detection, community detection [12, 13,14,15,16,17], and many more. Its incorporation with Ayurvedic practice has the potential to transform healthcare provision and personalized wellness offerings. The merger of machine learning and Ayurvedic practice will lead our culture to be more expandable all over the world through its truly good values and measures. In addition with the advances in machine learning we are seeing that hardware industry has made major jump by creating the digital devices such as smartwatches, smartphones, and fitness bands. These devices are very well equipped with sensors that can track a variety of health parameters such as step count, burned calories, sleep patterns, heart rate, and oxygen levels. This richness of their collected data is so much essential for accurate diagnosis and treatment planning. Recent research efforts had also concentrated on integrating this data with machine learning algorithms by opening the path for the creation of new frameworks that seamlessly combine these sensored data with Ayurvedic principles. Through this survey, we are aiming to lay the groundwork for our novel framework which also improve the limitations of smart data and Ayurveda to revolutionize healthcare systems and promote well-being all over the globe.

The organisation of this section is given below:

- First, the review dives deep into exploring research that predicts Ayurvedic doshas using a variety of machine learning and basic deep learning methods.
- Next, we will look at some studies that use smart data and machine learning to shed light on current trends and practices.
- And finally, because our main interest lies in understanding sleep quality, we will also review some research specifically focused on this aspect of health.

2.1.1 A Review on Human Constitution Classifications using ML and Basic DL Techniques

Majhi and colleagues [18] told us a machine-learning model merge with Ayurvedic principles for Parkinson's Disease (PD) prediction. Debnarayan Khatua and team [19] implemented a Dense Neural Network (DNN) to categorize different human constitutions. Researchers like T. Thanushree, K.G. Manjunath [20], Vani Rajasekar

[21], and H M Manjula along with AnandaRaj S P [22] employed a variety of machine learning techniques and ensemble methods to differentiate Ayurvedic constitution types (Prakriti) and tackle complex issues, reminiscent of seeking opinions from multiple doctors before deciding on a medical procedure. Vishu Madaan and Anjali Goyal [23] applied machine learning models and highlighted that further enhancements in boosting algorithms could advance Ayurvedic diagnostics. Tiwari P, Kutum R, Sethi T, and others [24] utilized machine learning to determine Ayurvedic constitution types (Prakriti) based on phenotypic characteristics. The discussed research demonstrates the effectiveness of supervised learning and some efforts to integrate basic deep learning models like ANN and DNN for diagnostic purposes. Murtaza M. Junaid Farooque [25] along with his collaborators adopted data mining techniques to classify human subjects by Ayurvedic Prakriti. Kaur Ranjit and Madaan Vishu, among others [26], introduced a fuzzy logic-based expert system crafted using MATLAB's Fuzzy Logic Toolbox, aimed at aiding physicians in delivering personalized and accurate treatments. Ayshree Ghorpade-Aher [27] proposed using a Feed Forward Neural Network to predict pulse variations due to Surya Namaskar across different Ayurvedic Prakriti and times of day (Prahar), using pulse data collected before and after the yoga sequence. Table 2.1 summarizes the primary outcomes of our literature review and provides a comparative evaluation of various models employed in Ayurvedic diagnostics.

TABLE 2.1: Comparative Evaluation of Various Models Employed in Ayurvedic Diagnostics

Year	Ref.	Model Used	Performance	Dataset	Pros	Cons
2023	[18]	LR, KNN, SVM, KSVM, NB, DT, RF, XGBoost	The top accuracy of 92.66% was achieved by Logistic Regression (LR).	Fox Insight dataset comprising 80,916 entries	Integrates Ayurvedic concepts with modern machine learning for Parkinson's Disease prediction. Achieves high accuracy (>92.5%) without reducing dimensionality.	Lack of application of more accurate deep learning techniques and omission of patients' medical history.
2023	[19]	Dense Neural Network	Precision: 0.94, Recall: 0.93, F-score: 0.93 for the western group;	Data from 233 extreme Prakriti individu	Automates Prakriti prediction and introduces transfer learning,	Constraints due to dataset quality and

			Precision: 1.0, Recall: 1.0, F-score: 1.0 for the northern group	als from western and northern regions	surpassing many advanced methods.	size; steep technological learning curve for Ayurvedic practitioners.
2022	[20]	SVM, XGB Classifier, Gaussian NB	High accuracy, recall, precision, and F-score with SVM and XGB Classifier	Data from 807 individuals	Useful for both well and unwell individuals; does not require Ayurvedic experts	Variability in prediction precision among individuals; authenticity issues.
2022	[21]	SVM, NB, DT, KNN, ANN, AdaBoost	AdaBoost scores 0.97 accuracy, 0.96 precision, 0.96 F-score, RMSE 0.64	Questionnaire data with 22 attributes from 484 healthy individuals	Employs ensemble learning to enhance model performance significantly.	Focused on Ayurvedic Dosha studies; quantitative reliability is still under development.
2021	[22]	XGBoost, AdaBoost, KNN, MLP, Random Forest, SGD	The highest accuracy of 97% was achieved with Random Forest and SGD	Questionnaire dataset	Enhances disease prediction relevant to modern lifestyles and allows treatment customization.	The dataset scale and the need for further validation are limitations.
2020	[23]	ANN, KNN,	Precision: 0.96, Recall:	Questionnaire	Superior precision,	Mainly focused

		SVM, NB, DT, Cat Boost	0.95, F-score: 0.95, Accuracy: 0.95	data from 807 individuals aged 20-60	recall, and accuracy; exceed traditional methods.	on constitution prediction; may lack broader Ayurvedic applications.
2017	[24]	LASSO, Elastic Net, Random Forest	Sensitivity: 88%-93%, Specificity: 90%-100%	Data from 147 healthy individuals	Precise predictions with fewer variables for Prakriti identification.	Limited sample size causes multicollinearity issues.
2017	[25]	Feed Forward Neural Network (FFNN)	Not specified; uses pulse waveform analysis for prediction	Pulse data from 30 individuals recorded over 4 days during Surya Namaskar	Merges traditional yoga with modern health monitoring; promotes a comprehensive health system.	Lacks detailed performance metrics and model validation, raising efficacy and applicability concerns.
2016	[26]	Decision Tree (J48), Naive Bayes, Logistic Regression, ANN (Multi-layer perceptron)	Best performance by ANN: Kappa: 0.1232, MAE: 0.3837, RMSE: 0.569	Data from 67 healthy subjects based on a 37-parameter questionnaire	Validates scalable machine learning application in Prakriti assessment.	Small sample size; focuses on a single aspect of Ayurvedic diagnosis; needs extensive

						validation for diverse populations.
2016	[27]	Fuzzy based System	Accuracy and veracity checked against Ayurvedic expert opinions.	Survey data from human subject studies	Successful use of fuzzy logic in traditional medicine, potentially improving personalized Ayurvedic treatment.	Complexity in design and need for expert interpretation; effectiveness dependent on accuracy.

2.1.2 Review of the Integration of Smart Data with Machine Learning Techniques for Healthcare

In recent years, the integration of machine learning (ML) and artificial intelligence (AI) in healthcare has significantly advanced, leading to the development of various innovative applications. In 2024, Damre et al. [28] introduced a Smart Healthcare Wearable Device for early disease detection, utilizing advanced machine learning algorithms to improve early diagnosis and patient monitoring. Moving slightly earlier, in 2023 [29,30], a development highlighted by an unnamed author involved a sleep quality prediction model that leverages deep learning to offer personalized health insights, reflecting an ongoing trend towards customized healthcare solutions. In the same year Seng et al. had also discussed the broadening scope of AI and ML in enhancing smart wearable devices by indicating that a rise in wearable technology capabilities across various sectors which also includes health and fitness of body.

In 2022, Sujith et al. [31] provided a comprehensive review of smart health monitoring systems that integrate AI and deep learning approaches by also emphasizing the good improvements in data analysis and personal care. This review underscores the major role of AI in transforming healthcare system through enhanced diagnostic accuracy and operational efficiency. The trend towards wearable health technology was also explored in 2021 [32] by authors who developed a smartwatch-based depression detection model using machine learning, as reported by an unnamed source. This model highlights the potential of wearables to contribute to mental health management by enabling continuous monitoring and early detection of depressive symptoms.

Earlier, in 2019 [33], Chandre et al. developed the HealthGuard which is a security framework for Smart Healthcare Systems (SHS) and which utilizes AI to secure data and ensure privacy in healthcare applications. This framework addresses critical concerns regarding the security and confidentiality of patient data in an increasingly digital healthcare environment. Collectively all these developments illustrate a major shift towards integrating ML techniques in healthcare majorly aiming to enhance patient outcomes, also to personalize treatment plans, and to secure sensitive health information. Table 2.2 provides the evaluation details of the integration of smart data with machine learning techniques in healthcare

TABLE 2.2: Merger of the Smart Data with Machine Learning Techniques in Healthcare

Year	Ref.	Model Used	Performance	Dataset	Pros	Cons
2024	[28]	Machine Learning Algorithms (unspecified types)	Promised High Accuracy in early disease detection	Data from integrated sensors on a wearable device	Early detection of diseases; personalized health recommendations; real-time data	Requires continuous data transmission; privacy concerns
2023	[29]	Deep Belief Network (DBN)	Accuracy: 97.50% Precision: 95.36% Recall: 94.87% F-Score: 95.02% G-Measure: 95.07%	The Sleep Study dataset from Kaggle consists of 400 samples	Improved sleep quality prediction; uses enhanced seagull optimization for tuning	Complexity of model and optimization process
2023	[30]	Machine Learning, Deep Learning, AI	Enhanced device functionalities	Various wearable device datasets	Expands the functionalities of wearable devices in various sectors.	Challenges in data processing and need for extensive training data.
2022	[31]	Deep Learning, AI, IoT	High accuracy in data analysis	Various smart health monitoring datasets	Improved accuracy and predictive capabilities in healthcare monitoring.	Complexity in integration and data privacy concerns.
2021	[32]	Logistic Regression, SVM, Decision Tree, K-NN, Naïve Bayes	89.603% for Ensemble Model 1 (K-NN, Logistic Regression, SVM) and 87.539% for	Kaggle dataset with 334 samples and 31 attributes	Utilizes readily available wearable data; provides	This may raise privacy concerns; limited by the quality of self-

			Ensemble Model 2 (Decision Tree, Naïve Bayes, SVM).		real-time monitoring	reported data
2019	[33]	ANN, Decision Tree, Random Forest, k-Nearest Neighbour	Accuracy: 91%, F1-score: 90%	Data from 8 different smart medical devices	Effective detection of malicious activities in SHS	Complex model setup, extensive data requirement

2.1.3 Review of Health Monitoring with the Help of Sleep Quality

By considering how important sleep is to preserving human health, a great deal of study has been done to carefully examine each person's unique sleep patterns in order to identify any irregularities as soon as possible. Furthermore, Ayurvedic Prakriti identification is a documented technique for identifying physiological abnormalities in humans. In addition to reviewing earlier studies that took place in the field of sleep analysis [34] of humans and also in Dosha prediction using a variety of machine learning (ML) techniques this will also look at approaches put forward by researchers to use data from smart devices for ML applications in healthcare industry.

In one study [35], researchers concentrated on EEG-based methods for assessing the quality of sleep, especially employing waking EEG signals to analyze sleep patterns from the night before. These studies often employ ML algorithms such as KNN, SVM, and GELM to classify sleep quality based on features extracted from EEG signals. Another study [36] introduced a two-step approach for determining sleep postures and stages. Initially, a stage detection algorithm utilizes variations in the FSR sensor to monitor user movement during sleep. Subsequently, a sleep position detection algorithm applies supervised learning to classify FSR signals into four positions: supine, prone, left lateral, and right lateral. In another piece of research [37], authors utilized a dataset from 54 patients undergoing polysomnography to label their findings, developing personalized ML models trained on individual-level actigraphy data to predict sleep-wake states and assess sleep characteristics. Further research [38] proposed using ML models to assess stress levels based on sleep data collected by SayoPillow, a multimodal dataset that monitors various physiological signals during sleep. An article [39] suggested a method to compute objective daily sleep habit scores from wearable device data, employing a stacking ensemble model for intermediate sleep states and logistic regression for distinguishing good from bad sleep. Research [40] highlighted how Vata, Pitta, and Kapha Doshas impact sleep patterns and quality, noting Vata's association with increased sleep onset latency, Pitta's correlation with sleep disturbances, and Kapha's tendency for deeper, longer sleep. These findings affirm the significance of Ayurveda in modern sleep research, suggesting that the integration of traditional knowledge with contemporary sleep analysis can deepen our understanding and enhance the customization of sleep health interventions. Figure 2.1 shows the researches took place in sleep analysis with machine learning.

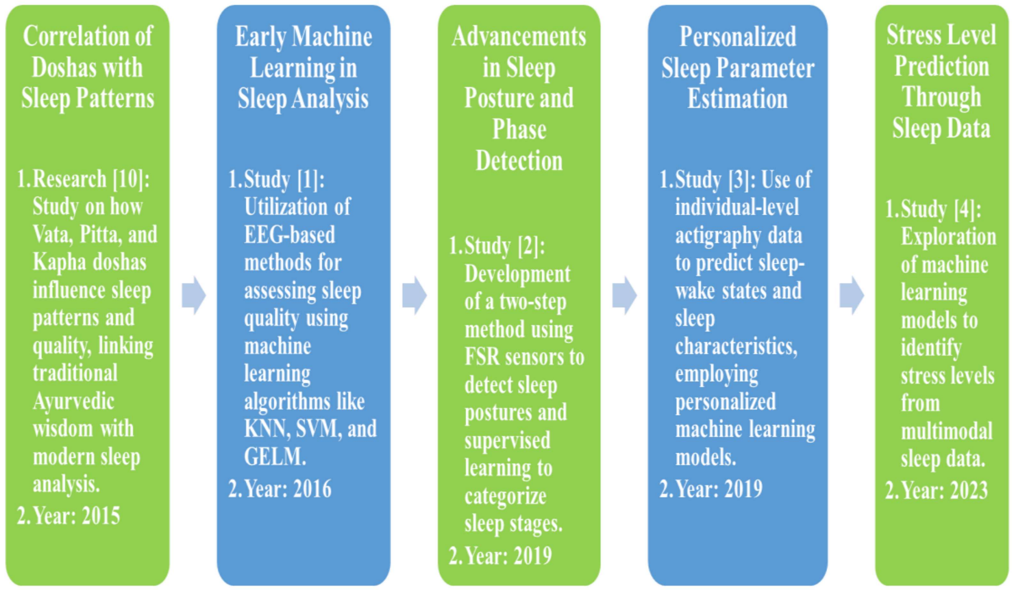


FIGURE 2.1: Advances in Sleep Analysis with Machine Learning

2.2 Algorithmic Frameworks in Ayurvedic Predictions and Smart Health Monitoring

The integration of recent smart health technologies with traditional Ayurvedic medicine creates new opportunities for enhancing treatment efficacy and diagnostic precision in the ever-evolving field of healthcare. This section examines several computational frameworks that have revolutionized smart health monitoring and Ayurvedic prediction. Every method, ranging from Fuzzy Based Systems to Logistic Regression, has advantages of its own when it comes to interpreting intricate health data collected from smart devices. These algorithms are combining the traditional knowledge with state-of-the-art technology which enable the accurate prediction of Ayurvedic Prakriti as well as it also provide complete health monitoring results. With the help of these frameworks Practitioners can more effectively give more information and also can obtain deeper insights into individual health trends by employing these machine-learning capabilities. In the other followed subsequent subsections a detailed analysis of each algorithm's and its workings has been done by also including their application and specific advantages in the context of integrating health metrics and Ayurvedic principles. Fig 2.2 shows a generalized framework for ML models for Ayurvedic predictions using smart device data.

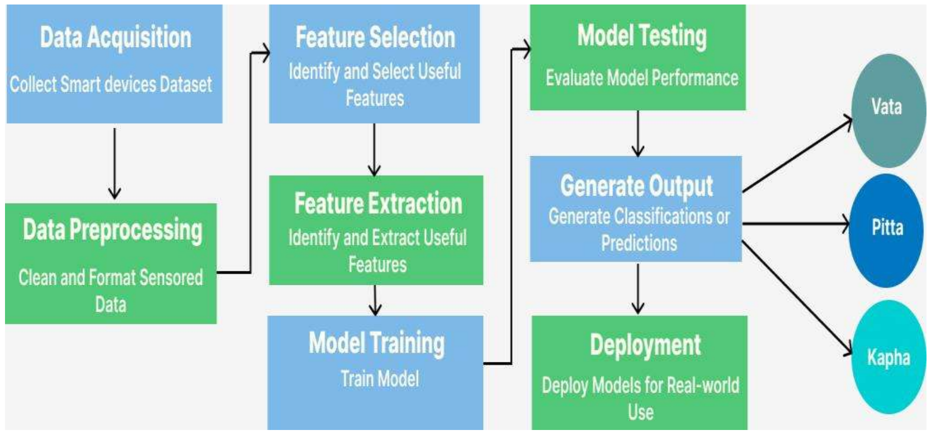


FIGURE 2.2: A Generalized Framework for Machine Learning models for predictions using smart device data

- **Logistic Regression:** Logistic Regression [41] is a statistical model which is used for the binary classification problems. It is very effective algorithm in predicting binary outcomes like the presence or absence of a individual Prakriti type based on health metrics dataset collected from smart devices. It also calculates probabilities using a logistic function which makes it more fast for situations where outputs are dichotomous.

- **Support Vactor Machines:** SVM [42] is a very powerful classification technique that works well on both linear and non-linear data structures. It is well used in Ayurvedic Prakriti prediction by finding the hyperplane that help in best separation of various classes of Prakriti types. It has been seen that it is utilized well with health metrics from smart devices. Also, SVM can classify complex health states with high accuracy.

- **Decision Tree (DT):** Decision Trees [43] are a non-linear predictive modeling tool that can be used very effectively in classifying human Prakriti based on physiological and psychological traits. They are very intuitive and easy to interpret, hence making them useful for decision-making in health applications on smart devices.

- **KNN:** KNN [44] is a simple and instance-based learning algorithm where the output is in the form of a class membership. It is used to predict an individual's Prakriti by finding the most similar cases within a small and also in a specified number of training samples (k) which is based on health metrics like pulse rate, blood pressure, etc., which are collected from smart devices.

- **Naïve Bayes:** This is a probabilistic classifier based on a theorem known as Bayes' theorem [45] which takes strong independence assumptions between the features. It is more suitable for large datasets and also it can be used to predict Prakriti types from health data collected by smart devices.

- **Random Forest:** Random Forest [46] is an ensemble learning method for classification and regression task that usually helps to build the multiple decision trees and then merge them to get a more accurate and stable output predictions. It is more sensible in the case of overfitting and is very effective in handling various health metrics collected from smart devices for accurate health monitoring and Prakriti classification.

- **Artificial Neural Networks (ANN):** ANNs [47] are inspired by the human brain, particularly neurons, and they consist of layers of interconnected neurons. They are highly flexible and powerful for modeling complex relationships in data which makes them ideal for interpreting complex patterns in the health metrics from smart devices to predict health and Prakriti types of humans.

- **Deep Belief Network (DBN):** DBN [48] is a part of deep neural networks that consists of multiple layers of graphical models that are having both directed and undirected edges in between. It is particularly good in feature extraction for the large amount of unstructured data.

- **XGBoost:** XGBoost [49] classifier stands for extreme Gradient Boosting and it is an implementation of gradient-boosted decision trees which is designed for its fast speed and performance. It is highly efficient at the classification tasks and also it can be used for fast Prakriti prediction by analyzing complex health metrics from smart devices.

- **AdaBoost:** AdaBoost [50], or Adaptive Boosting is an another ensemble learning technique based on decision trees that combines multiple weak classifiers to make a strong classifier. It is particularly very useful in improving the accuracy of prediction of various health metrics analysis from smart devices which in return ensures the enhanced diagnostic reliability.

- **Feed Forward Neural Network (FFNN):** FFNNs [51] are the simplest type of artificial neural network where the connections between the nodes do not form a cycle. This straightforward architecture is effective in various applications that include the categorization of Prakriti from structured health data collected via smart devices.

- **Fuzzy Based System:** Fuzzy systems [52] uses a fuzzy logic instead of traditional Boolean logic and is more suitable for handling t ambiguity in human health stages. This approach is effective in Ayurvedic applications where the diagnosis and treatment often involve interpreting complex and subjective health descriptions, which can be mirrored in data from smart health devices.

2.3 Evaluation Metrics in Ayurvedic Predictions and Smart Health Monitoring

It is important to provide an accurate evaluation in the application of machine learning algorithms particularly when we are integrating artificial intelligence with medical practices and here basically our traditional practice. This section namely "Evaluation Metrics in Ayurvedic Predictions and Smart Health Monitoring," helps us to understand and outlines the key metrics used to measure the effectiveness and reliability of ML models applied to health data from smart devices. Metrics such as accuracy, precision, recall, F1 score, sensitivity, and specificity [53,54,55,56,57,58] are used to provide useful insights into the model's performance in various dimensions which includes the evaluation from general correctness and its ability to distinguish between different health attributes and Prakriti types. The proper understanding of these metrics is so much essential for engineers and practitioners to refine these algorithms, and to enhance diagnostic accuracy, and also to ensure that the models are both clinically relevant and technically perfect. The subsequent discussion in Table 5 will explain each metric in detail by providing definitions and also by explaining their importance which are relevant for health analytics and personalized medicine based on Ayurvedic principles. Table 2.3 gives a brief introduction to different performance metrics used throughout by different authors.

TABLE 2.3: Different Evaluation metrics used throughout

Metric	Description
Accuracy	Measures the overall correctness of the model.
Precision	Measures the accuracy of positive predictions.
Recall	Measures the ability of a model to find all the relevant cases (True Positive).
F1 Score	Harmonic means of precision and recall.
Sensitivity	Same as recall, measures the proportion of actual positives correctly identified.
Specificity	Measures the proportion of actual negatives that are correctly identified.

2.4 Current Research Limitations

- Standardization of Data Collection Methods: The research consistently points to a need for uniform data collection methodologies within Ayurvedic practices. This standardization is critical to developing machine learning models that provide reliable and unbiased predictions. Improving the completeness and accuracy of the data collected will significantly bolster the effectiveness of these models.
- Enhancement of Dataset Quality: A notable limitation highlighted is the

lack of extensive and qualitative datasets in the domain of Ayurveda. There is a crucial need to cultivate large-scale datasets that encapsulate a broader spectrum of patient demographics and health conditions to enhance model training and accuracy.

- Expanding Machine Learning Methodologies: The current research primarily utilizes supervised learning approaches. There is an evident gap in the adoption of a wider array of machine learning strategies, particularly unsupervised learning, which could unearth new patterns and insights beneficial for Ayurvedic diagnostics. One expansion we have tried to cover with the help of applying a clustering algorithm to identify the human constitution.
- Robust Clinical Validation: For Establishing the clinical standards of machine learning applications in Ayurveda is very challenging procedure. Due to the subjective nature of Ayurvedic assessments throughout it is very complex task to standardize validation criteria. It is very essential to develop robust validation frameworks in order to clinically validate the procedure and to in crease the trust.
- Model Transparency and Understandability: There is need requirement for the development of machine learning models that are not only effective but also interpretable. The lack of transparency in current models, especially those based on deep learning lacks their trust of acceptance by healthcare practitioners. There is a need to develop more transparent models that would help in likely enhance trust and facilitate wider adoption in clinical settings.
- Technological Integration of Traditional Medicine with Smart Devices: As the risig of healthcare continues with the advancement of technology, there is a great opportunity to integrate traditional medicine systems such as Ayurveda with smart devices. These devices, which often include wearables like smartwatches and smartphones that are able to collect a wide range of health-related data in real time and also includes sleep patterns, heart rate, physical activity, and more. By using these types of data and with the help of machine learning models provide us with more personalized and accurate healthcare recommendations which are based on an individual's unique constitution as written in our traditional books and practiced by our doctors.

CHAPTER 3

METHODOLOGY

3.1 Experimental Setup

This study's motive is to categorize human Ayurvedic Prakriti (Vata, Pitta, and Kapha) using data captured and collected using smart devices that are used by individuals in their day-to-day life to make a way for personalized Ayurvedic healthcare and to standardize Ayurveda procedure which is at this is very subjective in nature. It analyzes sleep patterns collected from smartwatches and smartphones using an unsupervised learning algorithm to find similarities between individuals' Doshas.

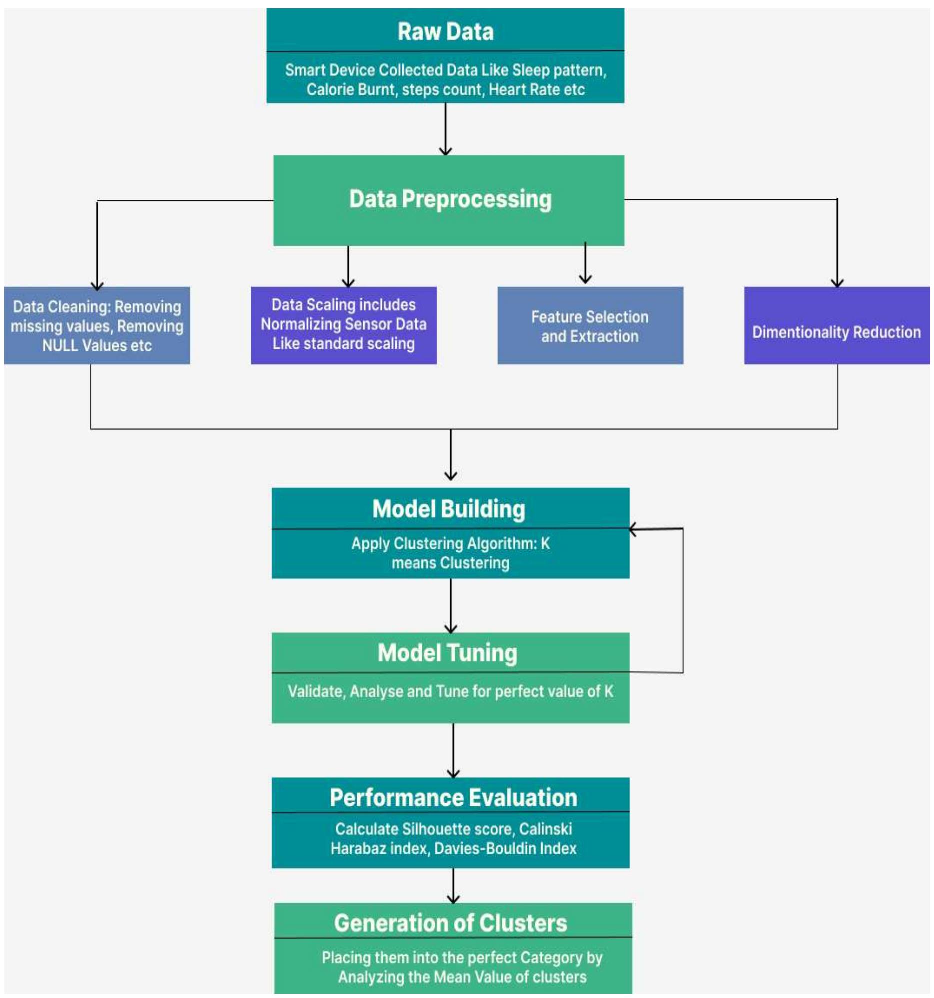


FIGURE 3.1: Generalized Framework for ASGF Model

Fig. 3.1 shows the generalized diagram of our proposed methodology which is all about firstly collecting data from smart devices. The diagram depicts that the first raw data from these devices is acquired or downloaded which contains attributes like sleep stages, heart rate, step counts, calories burnt, and so on. Secondly, there comes preprocessing where the data is cleaned, missing values are filed with the help of various techniques like up-filling, down-filling, mean fillings, and so on according to the characteristics of the data, and then the data is standardized using mean imputation method. After that, dimensionality reduction is taken into play to simplify the dataset and make it easier to deal with it. This step is important for improving computational efficiency and reducing noise in the data which ultimately enhances the accuracy of future analyses going to take place. As we employ unsupervised learning so clustering technique is used, where similar data points are grouped using an algorithm known as K-means Clustering [60]. Once that has been done, there is a fine-tuning step that takes place to make sure that the clustering of data is as accurate as possible. This step involves adjusting the parameters and evaluating different configurations to optimize the algorithms. Then comes the evaluation step which helps to measure the model performance and to analyze how well everything is working using different metrics in order to validate that everything is working fine or not. Finally, clusters are formed according to the similarity of datapoints. This proposed framework can make a way for personalized healthcare utilizing our cultural medical heritage.

3.2 Dataset Description and Preprocessing Overview

In this research, we utilized data from the SleepQual and B. Health datasets were collected from 24 university students over seven days and nights through their smartwatches and smartphones [61]. While the original datasets provided additional details on daily phone usage and physical activity, we chose to prioritize the examination of sleep data to gain a deeper insight into the specific correlation between sleep patterns and Ayurvedic Prakriti. Additionally, by directing all our efforts toward one aspect of health, we can effectively utilize our resources and ensure a thorough exploration of the link between sleep and Ayurvedic body types. Table 3.1 describes the attribute description of the dataset.

TABLE 3.1: Attribute Description

No.	Attribute	Description
1	Light Sleep Percentage	The proportion of total sleep duration is categorized as light sleep. Important for transitioning into deeper sleep stages and plays a role in brain detoxification and relaxation.
2	Deep Sleep Percentage	Percentage of the sleep cycle spent in deep sleep, crucial for physical recovery, immune system strengthening, and growth hormone release.

3	REM Sleep Percentage	Indicates the percentage of overall sleep time spent in REM sleep. Vital for emotional regulation, memory processing, and learning capabilities.
4	Label (Classification label)	Qualitative assessment of sleep quality, categorizing sleep into classes like 'good', 'poor', or 'excellent'.
5	Subject Identifier	A unique code or number is assigned to each participant to ensure anonymity and facilitate organized data analysis.
6	Day of Study	Marks the specific day on which the data was recorded, crucial for tracking patterns over the study duration.
7	Duration in Bed	Total time spent in bed during the recording session, including sleep time and periods of wakefulness, is usually measured in minutes.
8	In Bedtime (Time into bed)	The recorded time when the subject initially goes to bed is useful for analyzing sleep habits and latency.
9	Out of Bedtime (Time out of bed)	It is the time when the subject gets out of bed. It helps in the calculation of marking the end of the every sleep session and also it is useful for calculating total sleep time and efficiency of human.
10	Actual Sleep Duration	Actual amount of time spent asleep, excluding periods of wakefulness after initially falling asleep, measured in minutes.
11	Sleep Onset Latency	Measures the time it takes for the subject to fall asleep after getting into bed, an important metric for diagnosing sleep disorders like insomnia.
12	In Bed Awake Duration	Total time spent awake after initially falling asleep and before finally waking up, providing insights into sleep disturbances and overall sleep quality.
13	Sleep Efficiency	Calculated as the percentage of time in bed that was spent sleeping, a key metric for assessing the effectiveness of the sleep period.
14	Light Sleep Duration	Time spent in the light sleep phase, measured in minutes. Important for understanding the sleep architecture and its potential disruptions.
15	Deep Sleep Duration	Total duration spent in deep sleep during the sleep cycle quantified in minutes. Critical for physical health and recovery.
16	REM Sleep Duration	Duration of REM sleep throughout the night, measured in minutes, is essential for mental health and cognitive functions.
17	Awake Percentage	The percentage of the total time in bed that the subject remains awake. Helps to evaluate sleep efficiency and identify potential sleep issues.

For analyzing the data patterns effectively during the study we have utilized a rigorous preprocessing steps. In the Initial steps we have completed the cleaning of the dataset which is very helpful in removing any inconsistencies or errors present inside the dataset. Then after this we have performed exploratory data analysis to well understand the distribution and characteristics of the sleep-related attributes for selecting the useful features for the final model evaluation. This step involved like visualizing the data and identifying any outliers or unusual patterns.

Furthermore, we also employed statistical techniques like correlation analysis and regression modeling which helped us to derive the relationship between sleep patterns and Ayurvedic Prakriti. These analyses also helped us to quantify the strength and direction of the merger and also to identify potential attributes helpful in the identification of procedures of Prakriti types based on their sleep behaviors.

In the whole the analysis process we are very careful about the potential limitations like sample size limitations and data collection biases. We are able to addressed these challenges by employing sound statistical methods and also by conducting sensitivity analyses to assess the strength of our research findings.

The SleepQual dataset that we employed consist of a variety of sleep-related characteristics which includes duration, efficiency, onset delay, and various sleep stages as these are much more important for understanding the shading of sleep cycles and their relationship with Ayurvedic body types.

After cleaning and preparing the data, we proceeded with our study with the analysis. We used statistical techniques to find patterns and relationships within the data, focusing specifically on identifying how different sleep habits were aligning with different Ayurvedic Prakriti types. Fig 3.2. Represents the correlation matrix of various sleep attributes present inside the dataset. By systematically analyzing the data, we aimed to provide valuable insights by giving the relationship between sleep patterns and individual constitutions according to Ayurveda.

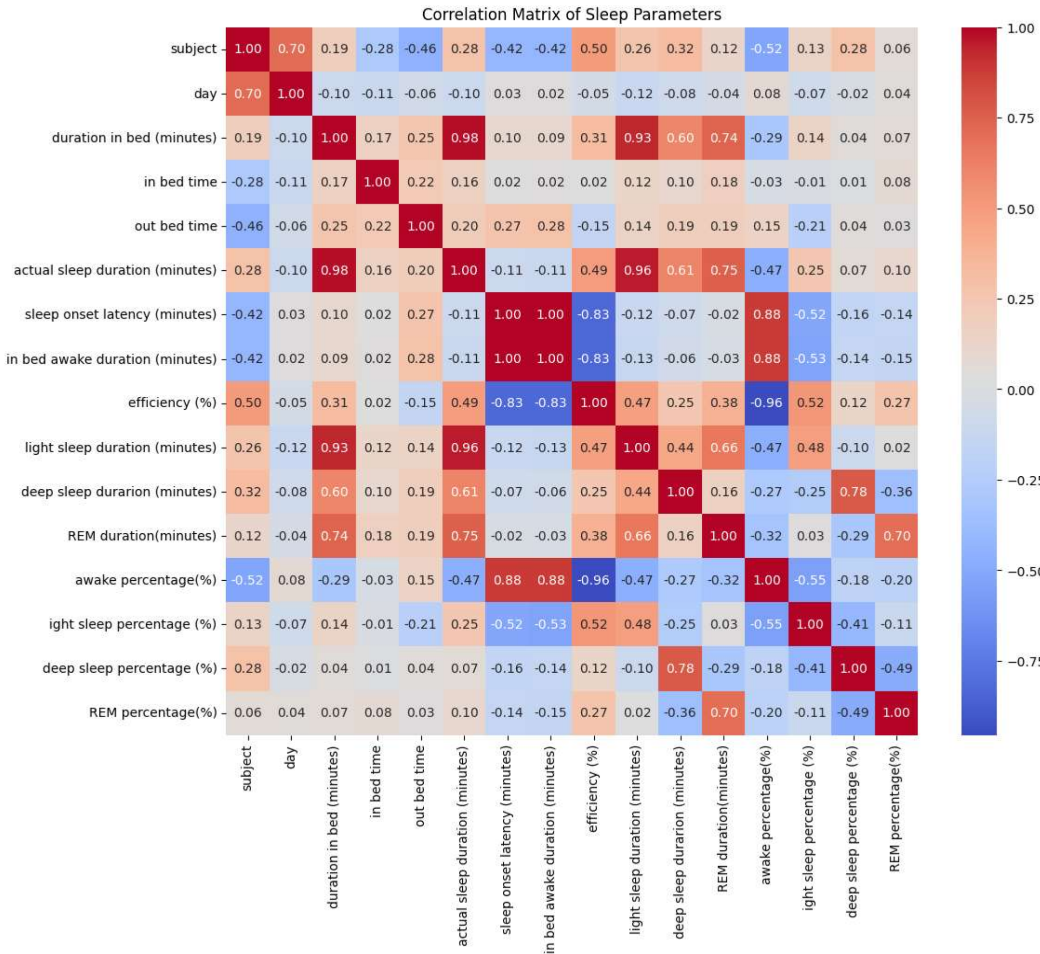


FIGURE 3.2: Correlation Matrix of Sleep Parameters

3.3 Data Analysis

The analysis of sleep patterns which we have performed, is to identify about how various sleep metrics are interconnected. The correlation matrix in the displayed figure provided insights into these relationships. Here is what we discovered: individuals who slept for longer tended to experience more light sleep, indicating a positive association between actual sleep duration and light sleep duration. Additionally, higher sleep efficiency correlated with quicker sleep onset and reduced time spent awake in bed. Moreover, there was a notable positive correlation between deep sleep duration and the percentage of time spent in deep sleep, which aligns with expectations. Interestingly, nights featuring more deep sleep tended to have less REM sleep, and vice versa. Fig. 3.3 shows the scatter pot between Actual Sleep Duration v/s Duration in Bed

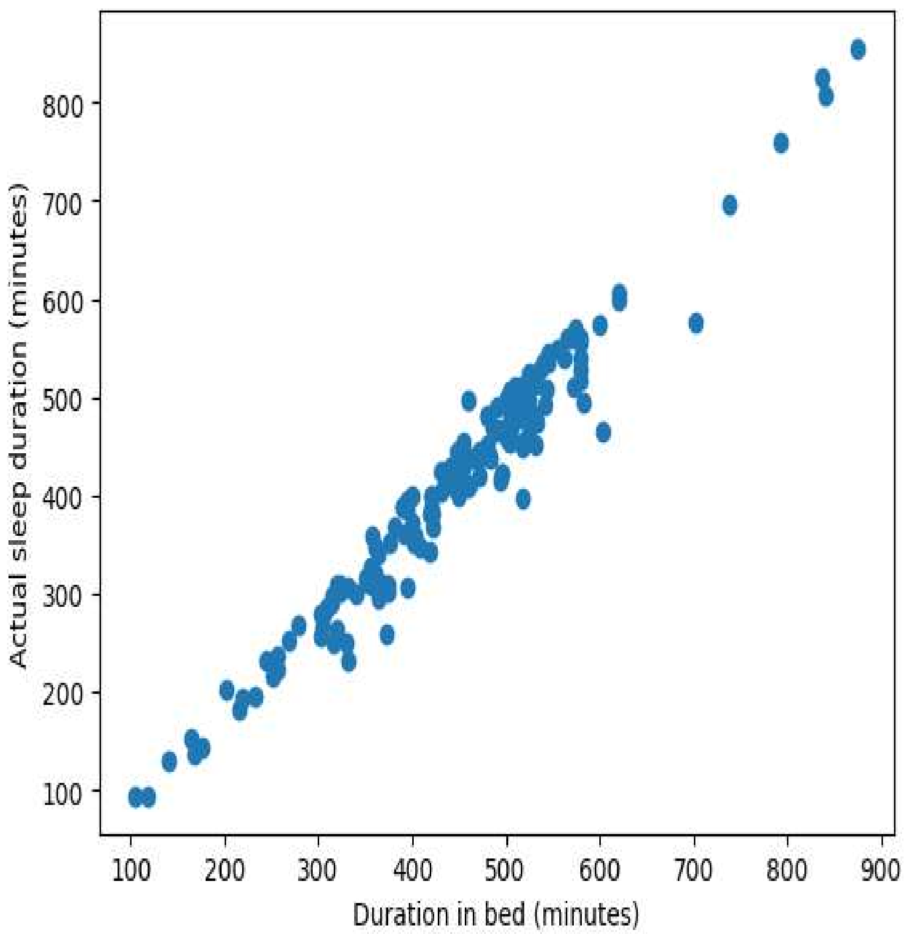


FIGURE 3.3: Scatter Plot between Actual Sleep Duration v/s Duration in Bed (min)

The dataset encompasses a diverse range of sleep durations, with an average nightly sleep duration of approximately 6.9 hours. On average, it takes individuals about 16 minutes to fall asleep, and most maintain a sleep efficiency of around 88-89%. Another graph illustrates a consistent pattern: individuals tend to sleep longer when they spend more time in bed.

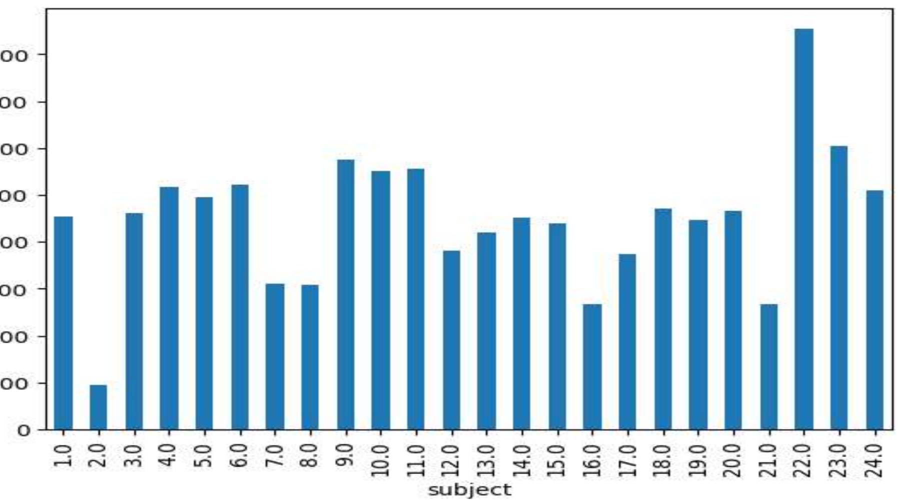


FIGURE 3.4: Bar Chart for Individual Sleep Time

Furthermore, we examined individual sleep times using a bar chart to discern any disparities between individuals who slept longer or shorter durations shown in Fig. 3.4. Lastly, a pie chart shown in Fig. 3.5 depicted the relative distribution of light sleep, deep sleep, and REM sleep stages. Light sleep was the most prevalent stage, followed by REM sleep, with deep sleep being the least experienced stage.

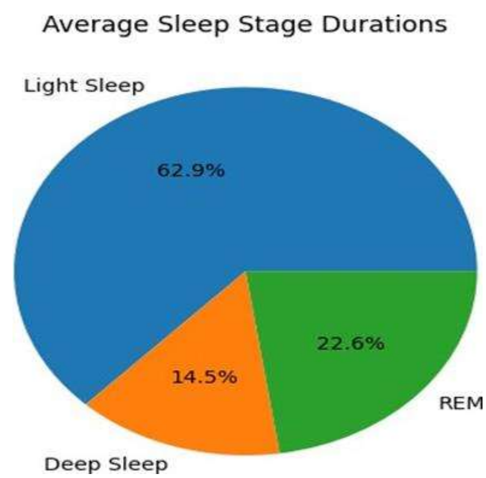


FIGURE 3.5: Pie Chart for Average Sleep Stage duration

3.4 Choice and Justification of Unsupervised Learning Techniques

The methodology of this thesis predominantly employs unsupervised learning techniques to analyze health data for Ayurvedic diagnostics. The decision to utilize unsupervised learning stems from its inherent capability to discover hidden patterns and structures from unlabeled data, which is crucial when dealing with Ayurvedic diagnostics, where explicit labels or categories are not always predefined or consistent. Among the various unsupervised techniques, K-means clustering was chosen due to its efficiency and effectiveness in identifying distinct, non-overlapping groups within data. This is very much relevant in the context of Ayurvedic diagnostics, where patient data must be segmented into specific categories representing different doshic imbalances (Vata, Pitta, Kapha) without prior labeling to help further in personalization. The algorithm which can handle large datasets and its soundness to produce stable clusters makes it suitable for this research which aims to map complex health metrics to traditional Ayurvedic concepts. Additionally, Principal Component Analysis (PCA) [61] was employed for dimensionality reduction before clustering. PCA helps in reducing the data into the components that are more informative and needed and that helps not only to get better visualization but also in enhancing the clustering process by reducing noise and redundancy in the data. This step is very critical in ensuring that the K-means algorithm on can focus on the most impactful features that are related to Ayurvedic principles and hence improving the interpretability and relevance for our findings and results.

3.5 Algorithmic Detail

3.4.1 Implementation of K-means Clustering

The main crux of our approach for identifying Ayurvedic Prakriti types from sleep data mainly involves the use of the K-means clustering algorithm. We used K-means in order to utilize its efficiency in handling large datasets and also due to its effectiveness in forming distinct, non-overlapping clusters based on various similarity measures.

3.4.2 Implementation Steps:

- Initialization: We started first by specifying the number of clusters, k, which, based on preliminary analysis and the number of primary Ayurvedic Prakriti types that is 3, three that is corresponding to Vata, Pitta, and Kapha prakriti. We choose the initial cluster of centroids randomly from the dataset.
- Assignment: Then each data point is assigned to the nearest cluster centroid with the help of the Euclidean distance metrics. This step also ensures that each data point is grouped according to the closest similarity in terms of sleep pattern features.
- Update: The centroids of the clusters are recalculated as the mean of all data points assigned to that cluster, which refines their position for better accuracy.
- Iteration: The assignment and update steps are repeated until the positions of the centroids stabilize, indicating convergence.

3.4.3 Custom Modifications for Ayurvedic Context:

- Feature Selection: Given the Ayurvedic focus, features related to sleep such as duration, quality, and efficiency are emphasized. This selective attention helps tailor the clustering process to aspects most relevant to Prakriti determination.
- Normalization: Data normalization is crucial for handling features that vary significantly in scale and distribution. By standardizing the data (subtracting the mean and dividing by the standard deviation), we ensure that each feature contributes equally to the distance calculations, preventing any one feature from dominating the algorithm due to scale differences.

3.4.4 Performance Metrics Employed

To evaluate the effectiveness of the clustering, we employed several metrics:

- Silhouette Score: This metric [62] measures the degree of separation between clusters. A higher silhouette score indicates that clusters are well-separated and cohesive, which is essential for ensuring that each Prakriti type is distinctly categorized.
- Davies-Bouldin Index: Used to assess the compactness and separation of the clusters. A lower Davies-Bouldin index [63] suggests that the clusters are well-separated and tightly grouped, which is desirable in a clustering scenario.
- Calinski-Harabasz Index: Another measure [64] used to evaluate the clusters based on their variance. A higher value indicates better-defined clusters.

3.5.5 Adaptations for Specific Research Needs:

To determine the optimal number of clusters, we implemented the elbow method, which involves plotting the sum of squared distances from each point to its assigned center as a function of the number of clusters. The point where the rate of decrease sharply shifts (the "elbow") [66] suggests the optimal number of clusters. This method is particularly useful in our study to confirm the preliminary assumption of three Prakriti types. Fig 3.6. Demonstrates the elbow method graph for finding optimal K value.

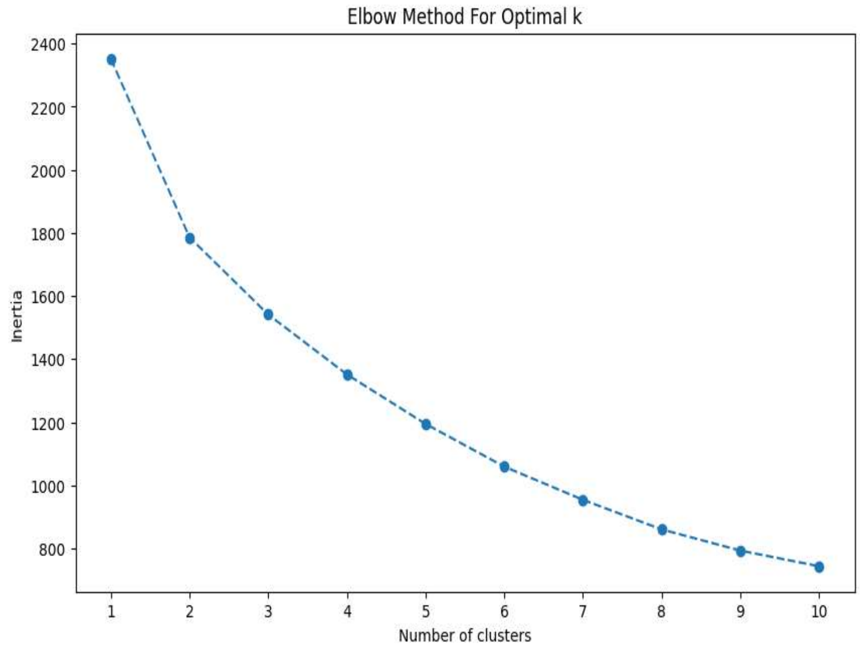


FIGURE 3.6: Elbow method for Optimal K

3.6 Pseudocode

Input: Smart device sleep data (e.g., duration, light sleep percentage, deep sleep percentage, REM percentage)
Output: Cluster assignments indicative of Ayurvedic Prakriti (Vata, Pitta, Kapha)

Initialization:
1. Specify the CLUSTERING_ALGO = K-means
2. N = Number of subjects in the dataset
3. k = Initial number of clusters (typically 3 for Vata, Pitta, Kapha)
4. I = Identity matrix for dimensionality reduction via PCA
5. Convergence criteria:
   - No significant change in intra-cluster distances, or
   - Optimal k value identified using the Elbow Method, or
   - Stability in cluster assignments across multiple iterations
6. Standardize features to have zero mean and unit variance for effective distance computation
7. min_points = Threshold for the minimal number of data points required per cluster; typically set based on data distribution and domain knowledge

Steps:
1. Load Data:
   - Import sleep pattern data from smart devices

- Parse data into usable format (e.g., matrices or data frames)

2. Data Preprocessing:
   - Clean data by handling missing values and anomalies
   - Scale data using standard normalization (Z-score normalization)
   - Select relevant features impacting sleep patterns
   - Apply PCA for dimensionality reduction to reduce computational complexity and enhance data interpretability

3. Model Building:
   - Initialize K-means clustering with specified k
   - Randomly assign initial cluster centers or use a smarter initialization method (e.g., k-means++)

4. Model Optimization:
   - Apply the Elbow Method to determine the optimal number of clusters:
     Calculate the total within-cluster sum of squares for each k
     Identify the k at which the rate of decrease sharply shifts (the elbow point)

5. Clustering Execution:
   - Run K-means clustering with the optimal k-determined
   - Iteratively update cluster centroids and reassign data points until convergence criteria are met

6. Evaluate Clustering:
   - Compute silhouette score for assessing how well-separated the clusters are
   - Calculate Davies-Bouldin and Calinski-Harabasz indices to evaluate cluster compactness and separation

7. Output Results:
   - Assign each subject to a cluster
   - Interpret and label clusters according to dominant Ayurvedic Prakriti characteristics observed in cluster members

End Algorithm

CHAPTER 4

RESULT AND ANALYSIS

4.1 Presentation of Clustering Results/ Cluster Analysis

The clustering process involved segmenting the sleep data into three distinct groups using the K-means algorithm. The initial step in the clustering involved the selection of the number of clusters (k=3) based on the Elbow Method, which suggested a significant reduction in within-cluster variance at this point. Scatter plots and heatmaps were utilized to visualize the distribution of data points within each cluster, illustrating the cohesive and distinct nature of each group. Cluster Interpretation: Each cluster was analyzed and interpreted in the context of Ayurvedic Prakriti as shown in Table 4.1. The clusters corresponded effectively to the expected characteristics of Vata, Pitta, and Kapha:

TABLE 4.1: Cluster Assignments Characteristics

Cluster	**Mean Properties**	**Ayurvedic Prakriti**	**Characteristics Alignment**
0	Duration in Bed (min): 295.89 Actual Sleep Duration (min): 268.25 Efficiency (%): 90.32 Light Sleep Duration (min): 168.69 Deep Sleep Duration (min): 42.05 REM Sleep Duration (min): 58.76	Vata	Cluster 1 (Vata) was characterized by irregular and shorter sleep durations with quick transitions between sleep stages.
1	Duration in Bed (min): 513.45 Actual Sleep Duration (min): 493.51 Efficiency (%): 95.90 Light Sleep Duration (min): 313.68 Deep Sleep Duration (min): 69.98 REM Sleep Duration (min): 110.94	Kapha	Cluster 2 (Pitta) showed moderate sleep durations with instances of sleep disturbances possibly due to an overactive mind.

2	Duration in Bed (min): 463.11 Actual Sleep Duration (min): 391.69 Efficiency (%): 84.38 Light Sleep Duration (min): 241.48 Deep Sleep Duration (min): 57.46 REM Sleep Duration (min): 94.65	Pitta	Cluster 3 (Kapha) displayed patterns of long, deep sleep, indicating a stable and heavy sleep profile.

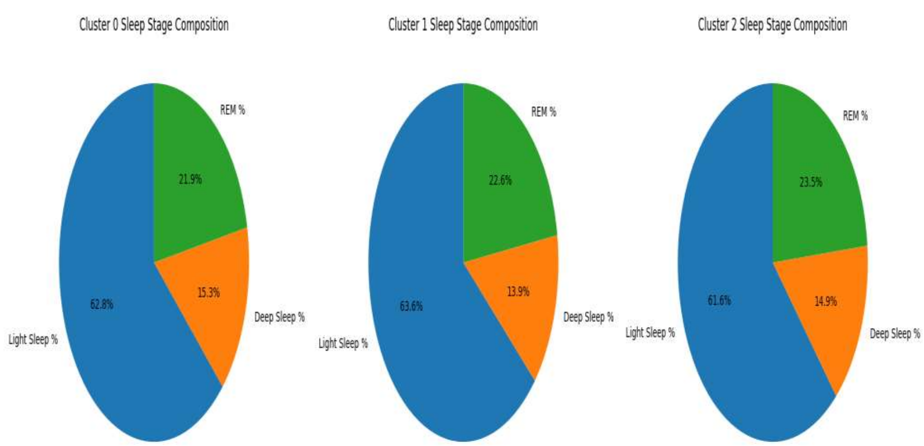


FIGURE 4.1: Pie chart for clusters Sleep State Composition

The provided pie charts in Fig. 4.1 visually represent the composition of sleep stages for three distinct clusters identified through K-means clustering of sleep pattern data, a key analytical method outlined in the thesis. Each chart corresponds to a cluster—0, 1, and 2—and illustrates the percentages of light sleep, deep sleep, and REM sleep, aligning with the typical sleep characteristics associated with the Ayurvedic Prakriti types. For example, Cluster 0, which is predominantly Vata, shows a high percentage of light sleep at 62.8% and lesser amounts of deep sleep and REM sleep, reflecting the irregular and lighter sleep patterns typical of the Vata constitution. Similarly, Cluster 1 and Cluster 2, representing Kapha and Pitta respectively, demonstrate their unique distributions of sleep stages that corroborate with their Ayurvedic descriptions. Kapha is known for its deep and reviving sleep characteristics. Also, Pitta-overloaded humans are generally associated with moderate sleep durations which are disturbed due to their overactive mind, and their patterns are captured with their corresponding sleep stages with the help of smart devices. The pie charts are so important for visualizing the dataset insights which are taken from the implementation by providing an insightful understanding of how different Prakriti types respond during sleep stages with the help of unsupervised machine-learning techniques. Fig. 4.2 represents a scatter plot that shows the results of a two-dimensional Principal Component Analysis (PCA) for clustering sleep data attributes into three distinct clusters and each is marked by different colors which are red, blue, and green. The x-axis is labeled "Principal Component 1" and the y-axis is labeled "Principal Component 2," which shows the distribution of the data points along these two principal components as it is difficult to analyze without the use of PCA.

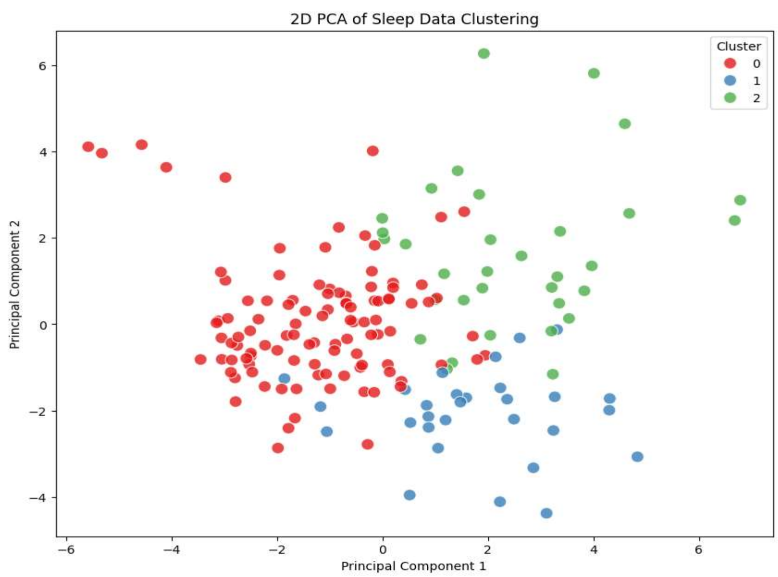


FIGURE 4.2: Clustering of Data

4.2 Performance Analysis

The effectiveness of the clustering was quantitatively evaluated using several performance metrics:

- Silhouette Score: The silhouette score of 0.2296 indicated a reasonable separation between clusters, albeit suggesting room for further optimization.
- Davies-Bouldin Index: A score of 1.5345 pointed to a decent level of cluster compactness and separation, affirming the distinctiveness of

each cluster.

- Calinski-Harabasz Index: A score of 43.3666 in the clusters demonstrated a good level of tightness and separation that also validated the effectiveness of the clustering data points.

4.3 Model Limitations

1. **Data Representation Bias:**
The datasets used in this study were primarily sourced from specific demographic groups which in return may not comprehensively represent the broader population. These biases could majorly affected the model's ability to accurately generalize its predictions across different ethnicities, age groups, and geographical locations. Additionally, the data might not fully capture the diversity of symptoms or presentations of Ayurvedic doshas that can vary significantly across individuals.

2. **Historical Bias:**
The training data reflects historical patient records and also may incorporate inherent biases in diagnostic practices or in the treatment approaches that were not dealt with at the time of data collection. This can lead the machine learning models to increase these biases unless specifically addressed through techniques like bias correction or re-sampling. This can be handled with the help of mapping correctly the history of patients with there other data.

3. **Risks of Overfitting:**
Machine learning models particularly those that involve complex structures like neural networks are so much prone to overfitting especially when trained on limited or highly specific data sets. Overfitting can generally occur when a model learns the detail and noise in the training data to the extent that it negatively impacts the performance of the model on new data. This is often seen by showing excellent performance on the training dataset but poor generalization to new or unseen data. So to solve this problem techniques such as cross-validation, regularization, and pruning are utilized. However, the risk of overfitting cannot be eliminated completely.

4. **Model Transferability:**
The generalizability of the machine learning models developed in this study is limited by the specific conditions under which the models were trained. These conditions include the nature of the data, the specific features used, and the configuration of the machine learning algorithms. So therefore in such a case, the models might not perform with the same accuracy when deployed in different Ayurvedic diagnostic settings or with data collected under different conditions.

5. **Scalability to Other Health Conditions:**
   The models which we have been using are successful in giving tailored responses in diagnosing Ayurvedic doshas but the main issue is that their scalability to other health conditions or their utility in predicting other types of medical outcomes remains untested. Therefore, expanding these models to cover a broader range of conditions would require not only more diverse data but also adjustments to the algorithms to accommodate different types of diagnostic criteria is very much needed.

CHAPTER 5

CONCLUSION AND FUTURE WORK

This study has successfully done the integration of machine learning and Ayurvedic practices through the analysis of health data collected from smart monitoring devices like smartwatches. It utilizes unsupervised learning algorithms in which we specifically used the K-means clustering algorithm with the help of which we have identified patterns that are more directly related to the practice and it also helped in showcasing the potential of modern technology to enhance the accuracy and applicability of ancient medical wisdom. The final findings of our study reveal that machine learning can take a characteristically advantage from the features of Vata, Pitta, and Kapha doshas through physiological and behavioral raw data derived from smart technology which as a result is well aligned with Ayurvedic principles. However, our research highlights the need for further validation and expansion due to a small dataset which can somehow give us biased results. Future work will aim to include a more diverse demographic to enhance the generalizability of our findings. There is also an opportunity to explore additional machine learning models, such as deep learning and reinforcement learning, to refine the predictions and adapt the diagnostic processes dynamically based on continuous health data. Collaborative efforts between AI engineers and Ayurvedic practitioners will be crucial in advancing this collaboration that ensures that the models that are developed are not only technically proficient but also practically applicable in real-world medical practice. Moreover, ethical considerations, particularly regarding data privacy and the security of health information have to be dealt with as we move forward towards the more complex model structure. It could be done by implementing robust privacy measures and by ensuring transparent data handling practices that will be essential to maintaining trust and integrity in the use of AI in healthcare. In conclusion, this research clears the way for an application approach to healthcare where majorly traditional Ayurvedic practices are enhanced by the precision and analytical power of machine learning and deep learning. As we move forward to continue expanding and refining these technologies then at the moment the potential for personalized and preventive healthcare becomes increasingly important by promising significant improvements in patient outcomes and the overall effectiveness of Ayurvedic diagnostics.

REFERENCES

1. Mukherjee, P. K., Harwansh, R. K., Bahadur, S., Banerjee, S., Kar, A., Chanda, J., ... & Katiyar, C. K. (2017). Development of Ayurveda–tradition to trend. Journal of Ethnopharmacology, 197, 10-24.
2. Hankey, A. (2010). Establishing the Scientific Validity of Tridosha Part 1: Doshas, Subdoshas: and: Dosha Prakritis. Ancient science of life, 29(3), 6-18.
3. Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. Science, 349(6245), 255-260.
4. Zhang, C., Liu, C., Zhang, X., & Almpanidis, G. (2017). An up-to-date comparison of state-of-the-art classification algorithms. Expert Systems with Applications, 82, 128-150.
5. Huang, J. C., Ko, K. M., Shu, M. H., & Hsu, B. M. (2020). Application and comparison of several machine learning algorithms and their integration models in regression problems. Neural Computing and Applications, 32(10), 5461-5469.
6. Abiodun, O. I., Jantan, A., Omolara, A. E., Dada, K. V., Mohamed, N. A., & Arshad, H. (2018). State-of-the-art in artificial neural network applications: A survey. Heliyon, 4(11).
7. Xu, R., & Wunsch, D. (2005). Survey of clustering algorithms. IEEE Transactions on Neural Networks, 16(3), 645-678.
8. Redfern, J. (2017). Smart health and innovation: facilitating health-related behavior change. Proceedings of the Nutrition Society, 76(3), 328-332.
9. Saravanan, R., & Sujatha, P. (2018, June). A state of art techniques on machine learning algorithms: a perspective of supervised learning approaches in data classification. In 2018 Second international conference on intelligent computing and control systems (ICICCS) (pp. 945-949). IEEE.
10. Naeem, S., Ali, A., Anam, S., & Ahmed, M. M. (2023). An unsupervised machine learning algorithms: Comprehensive review. International Journal of Computing and Digital Systems.
11. Holmes, J., Sacchi, L., & Bellazzi, R. (2004). Artificial intelligence in medicine. Ann R Coll Surg Engl, 86, 334-8.
12. A. Kumar, N. Aggarwal, S. Kumar, "SIRA: a model for propagation and rumor control with epidemic spreading and immunization for healthcare 5.0" Soft Computing, 27(7), 4307-4320 (2023)
13. A. Kumar, D.K. Jain, A. Mallik, S. Kumar, "Modified node2vec and attention based fusion framework for next POI recommendation". Information Fusion, 101, 101998 (2024)
14. S. Kumar, "Negative Stances Detection from Multilingual Data Streams in Low-Resource Languages on Social Media Using BERT and CNN-Based Transfer Learning Model", ACM Transactions on Asian and Low-Resource Language Information Processing, 23(1), 1–18 (2024)
15. S. Kumar, A. Kumar, A. Mallik, R. R. Singh, "OptNet-Fake: Fake News Detection in Socio-Cyber Platforms Using Grasshopper Optimization and Deep Neural Network" IEEE Transactions on Computational Social Systems. (2023).
16. A. Mallik, S. Kumar, "Word2Vec and LSTM based deep learning technique for context-free fake news detection". Multimedia Tools and Applications, 83, pages 919–940, (2024)
17. S. Kumar, A Mallik, S.S Sengar, "Community detection in complex networks using stacked autoencoders and crow search algorithm". The Journal of Supercomputing, 79(3), 3329-3356 (2023)
18. 20. Majhi, Vinayak, Bishnu Choudhury, Goutam Saha, and Sudip Paul.

"Development of a machine learning-based Parkinson's disease prediction system through Ayurvedic dosha analysis." *International Journal of Ayurvedic Medicine* 14, no. 1 (2023): 180-189.

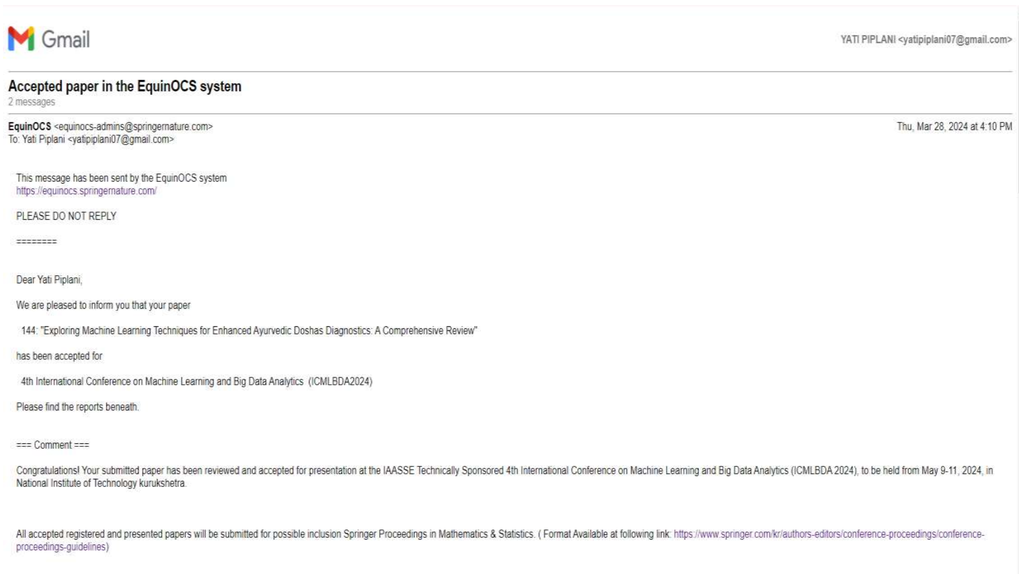
19. 21. Khatua, Debnarayan, Arif Ahmed Sekh, Rintu Kutum, Mitali Mukherji, Bhavana Prasher, and Samarjit Kar. "Classification of Ayurveda constitution types: a deep learning approach." *Soft Computing* 27, no. 9 (2023): 5309-5317.
20. 22. Thanushree, T., and K. G. Manjunath. "Ayurvedic Medicine-based Automated Machine Learning Approach for Determining the Nature of the Human Body." *Journal of Pharmaceutical Negative Results* (2022): 1161-1165.
21. 23. Rajasekar, V., S. Krishnamoorthi, M. Saraˇcevic, D. Pepic, M. Zajmovic, and H. Zogic. "Ensemble machine learning methods to predict the balancing of ayurvedic constituents in the human body: ensemble machine learning methods to predict, Comput. Sci. 23 (1)(2022)."
22. 24. Manjula, H. M., and S. P. AnandaRaj. "Ayurvedic Diagnosis using Machine Learning Techniques to examine the diseases by extracting the data stored in AyurDataMart." In *2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)*, pp. 239-244. IEEE, 2021.
23. 25. Madaan, Vishu, and Anjali Goyal. "Predicting ayurveda-based constituent balancing in human body using machine learning methods." *IEEE Access* 8 (2020): 65060-65070.
24. 26. Tiwari, Pradeep, Rintu Kutum, Tavpritesh Sethi, Ankita Shrivastava, Bhushan Girase, Shilpi Aggarwal, Rutuja Patil et al. "Recapitulation of Ayurveda constitution types by machine learning of phenotypic traits." *PloS one* 12, no. 10 (2017): e0185380.
25. 27. Ghorpade-Aher, Jayshree, Abhishek Patil, Yudhishthir Deshpande, Sumant Gaopande, and Eeshan Phatak. "A proposed framework for prediction of pulse, based on the effect of surya namaskar on different prakruti at different prahars of the day." In *2017 International Conference on Big Data Analytics and Computational Intelligence (ICBDAC)*, pp. 11-16. IEEE, 2017.
26. 28. Farooque, Murtaza M. Junaid, Mohammed Aref, Mohammed Imran Khan, and Shareque Mohammed. "Data Mining Application in Classification Scheme of Human Subjects According to Ayurvedic Prakruti â€"Temperament." *Indian Journal of Science and Technology* (2016).
27. 29. Ranjit, Kaur, Madaan Vishu, Agrawal Prateek, Singh Sanjay Kumar, and Kaur Amandeep. "Fuzzy expert system for identifying the physical constituents of a human body." *Indian Journal of Science and Technology* (2016).
28. Hamza, M. A., Abdalla Hashim, A. H., Alsolai, H., Gaddah, A., Othman, M., Yaseen, I., ... & Zamani, A. S. (2023). Wearables-assisted smart health monitoring for sleep quality prediction using optimal deep learning. Sustainability, 15(2), 1084.
29. Motwani, A., Shukla, P. K., & Pawar, M. (2022). Ubiquitous and smart healthcare monitoring frameworks based on machine learning: A comprehensive review. Artificial Intelligence in Medicine, 134, 102431.
30. Sujith, A. V. L. N., Sajja, G. S., Mahalakshmi, V., Nuhmani, S., & Prasanalakshmi, B. (2022). Systematic review of smart health monitoring using deep learning and Artificial intelligence. Neuroscience Informatics, 2(3), 100028.
31. Chiang, P. H., Wong, M., & Dey, S. (2021). Using wearables and machine learning to enable personalized lifestyle recommendations to improve blood pressure. IEEE Journal of Translational Engineering in Health and Medicine, 9, 1-13.

32. Kumar, P., Chauhan, R., Stephan, T., Shankar, A., & Thakur, S. (2021, January). A machine learning implementation for mental health care. Application: smart watch for depression detection. In 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence) (pp. 568-574). IEEE.
33. Newaz, A. I., Sikder, A. K., Rahman, M. A., & Uluagac, A. S. (2019, October). Healthguard: A machine learning-based security framework for smart healthcare systems. In 2019 sixth international conference on social networks analysis, management and security (SNAMS) (pp. 389-396). IEEE.
34. Satapathy, S., Loganathan, D., Kondaveeti, H. K., & Rath, R. (2021). Performance analysis of machine learning algorithms on automated sleep staging feature sets. CAAI Transactions on Intelligence Technology, 6(2), 155-174.
35. Wang, L. L., Zheng, W. L., Ma, H. W., & Lu, B. L. (2016, July). Measuring sleep quality from EEG with machine learning approaches. In 2016 International Joint Conference on Neural Networks (IJCNN) (pp. 905-912). IEEE.
36. Satapathy, S., Loganathan, D., Kondaveeti, H. K., & Rath, R. (2021). Performance analysis of machine learning algorithms on automated sleep staging feature sets. CAAI Transactions on Intelligence Technology, 6(2), 155-174.
37. Crivello, A., Palumbo, F., Barsocchi, P., La Rosa, D., Scarselli, F., & Bianchini, M. (2019). Understanding human sleep behaviour by machine learning. Cognitive infocommunications, theory and applications, 227-252.
38. Khademi, A., El-Manzalawy, Y., Master, L., Buxton, O. M., & Honavar, V. G. (2019). Personalized sleep parameters estimation from actigraphy: a machine learning approach. Nature and Science of Sleep, 387-399.
39. Rahman, M. M., Islam, A. M., Miah, J., Ahmad, S., & Mamun, M. (2023, June). sleepWell: Stress Level Prediction Through Sleep Data. Are You Stressed?. In 2023 IEEE World AI IoT Congress (AIIoT) (pp. 0229-0235). IEEE.
40. Kim, J., & Park, M. (2023). A Study on ML-Based Sleep Score Model Using Lifelog Data. Applied Sciences, 13(2), 1043.
41. Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). Applied logistic regression. John Wiley & Sons.
42. Schuldt, C., Laptev, I., & Caputo, B. (2004, August). Recognizing human actions: a local SVM approach. In P
43. roceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004. (Vol. 3, pp. 32-36). IEEE.
44. Priyam, A., Abhijeeta, G. R., Rathee, A., & Srivastava, S. (2013). Comparative analysis of decision tree classification
45. algorithms. International Journal of current engineering and technology, 3(2), 334-337.
46. Kramer, O., & Kramer, O. (2013). K-nearest neighbors. Dimensionality reduction with unsupervised nearest neighbors, 13-23.
47. Webb, G. I., Keogh, E., & Miikkulainen, R. (2010). Naïve Bayes. Encyclopedia of machine learning, 15(1), 713-714.
48. Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. ISPRS journal of photogrammetry and remote sensing, 114, 24-31.
49. Zupan, J. (1994). Introduction to artificial neural network (ANN) methods: what they are and how to use them. Acta Chimica Slovenica, 41(3), 327.
50. Hinton, G. E. (2009). Deep belief networks. Scholarpedia, 4(5), 5947.
51. Ramraj, S., Uzir, N., Sunil, R., & Banerjee, S. (2016). Experimenting XGBoost algorithm for prediction and classification of different datasets. *International Journal of Control Theory and Applications*, *9*(40), 651-662.

52. Ying, C., Qi-Guang, M., Jia-Chen, L., & Lin, G. (2013). Advance and prospects of AdaBoost algorithm. Acta Automatica Sinica, 39(6), 745-758.
53. Bebis, G., & Georgiopoulos, M. (1994). Feed-forward neural networks. Ieee Potentials, 13(4), 27-31.
54. de Campos Souza, P. V. (2020). Fuzzy neural networks and neuro-fuzzy networks: A review the main techniques and applications used in the literature. Applied soft computing, 92, 106275.
55. Yin, M., Wortman Vaughan, J., & Wallach, H. (2019, May). Understanding the effect of accuracy on trust in machine learning models. In *Proceedings of the 2019 chi conference on human factors in computing systems* (pp. 1-12).
56. Michaud, E. J., Liu, Z., & Tegmark, M. (2023). Precision machine learning. Entropy, 25(1), 175.
57. Davis, J., & Goadrich, M. (2006, June). The relationship between Precision-Recall and ROC curves. In Proceedings of the 23rd international conference on Machine learning (pp. 233-240).
58. Lipton, Z. C., Elkan, C., & Narayanaswamy, B. (2014). Thresholding classifiers to maximize F1 score. arXiv preprint arXiv:1402.1892.
59. Menon, P., Geevasinga, N., Yiannikas, C., Howells, J., Kiernan, M. C., & Vucic, S. (2015). Sensitivity and specificity of threshold tracking transcranial magnetic stimulation for diagnosis of amyotrophic lateral sclerosis: a prospective study. The Lancet Neurology, 14(5), 478-484.
60. Kodinariya, T. M., & Makwana, P. R. (2013). Review on determining number of Cluster in K-Means Clustering. International Journal, 1(6), 90-95.
61. Arora, A., Chakraborty, P., & Bhatia, M. P. S. (2022, January). SleepQual and B. health: Smartwatch and Smartphone based behavioral Datasets of youth. In 2022 12th International Conference on Cloud Computing, Data Science & Engineering (Confluence) (pp. 340-344). IEEE.
62. Maćkiewicz, A., & Ratajczak, W. (1993). Principal components analysis (PCA). Computers & Geosciences, 19(3), 303-342.
63. Shutaywi, M., & Kachouie, N. N. (2021). Silhouette analysis for performance evaluation in machine learning with applications to clustering. Entropy, 23(6), 759.
64. Xiao, J., Lu, J., & Li, X. (2017). Davies Bouldin Index based hierarchical initialization K-means. Intelligent Data Analysis, 21(6), 1327-1338.
65. Łukasik, S., Kowalski, P. A., Charytanowicz, M., & Kulczycki, P. (2016, July). Clustering using flower pollination algorithm and Calinski-Harabasz index. In 2016 IEEE congress on evolutionary computation (CEC) (pp. 2724-2728). IEEE.
66. Bholowalia, P., & Kumar, A. (2014). EBK-means: A clustering technique based on elbow method and k-means in WSN. International Journal of Computer Applications, 105(9).

PUBLICATIONS

Yati Piplani, Pawan Singh Mehra, "Exploring Machine Learning Techniques for Enhanced Ayurvedic Doshas Diagnostics: A Comprehensive Review "The paper has been accepted in 4th International Conference on Machine Learning and Big Data Analytics (ICMLBDA2024)
May 9-11, 2024, Indexed by Scopus. Paper Id: 144.



Gmail

YATI PIPLANI <yatipiplani07@gmail.com>

Accepted paper in the EquinOCS system
2 messages

EquinOCS <equinocs-admins@springernature.com> Thu, Mar 28, 2024 at 4:10 PM
To: Yati Piplani <yatipiplani07@gmail.com>

This message has been sent by the EquinOCS system
https://equinocs.springernature.com/

PLEASE DO NOT REPLY

========

Dear Yati Piplani,

We are pleased to inform you that your paper

144: "Exploring Machine Learning Techniques for Enhanced Ayurvedic Doshas Diagnostics: A Comprehensive Review"

has been accepted for

4th International Conference on Machine Learning and Big Data Analytics (ICMLBDA2024)

Please find the reports beneath.

=== Comment ===

Congratulations! Your submitted paper has been reviewed and accepted for presentation at the IAASSE Technically Sponsored 4th International Conference on Machine Learning and Big Data Analytics (ICMLBDA 2024), to be held from May 9-11, 2024, in National Institute of Technology kurukshetra.

All accepted registered and presented papers will be submitted for possible inclusion Springer Proceedings in Mathematics & Statistics. (Format Available at following link: https://www.springer.com/kr/authors-editors/conference-proceedings/conference-proceedings-guidelines)

Yati Piplani, Pawan Singh Mehra, "An Unsupervised Learning Model for Ayurvedic Prakriti Determination Using Smart Device Data" The paper has been accepted in at OPJU International Technology Conference (OTCON 3.0) on Smart Computing for Innovation and Advancement in Industry 4.0 June 5-7 2024, Indexed by Scopus in IEEE. Paper Id: 820.

OTCON 3.0: Acceptance of Paper ID-820 at OPJU International Technology Conference (OTCON 3.0) on Smart Computing for Innovation and Advancement in Industry 4.0 Inbox ×

Microsoft CMT <email@msr-cmt.org>
to me

Thu, Apr 25, 10:09 AM

Dear Yati Piplani,
Paper ID: "820"
Title: "An Unsupervised Learning Model for Ayurvedic Prakriti Determination Using Smart Device Data"

Congratulations!!!
We are delighted to inform you that above mentioned paper has been ACCEPTED for oral presentation at the OPJU International Technology Conference (OTCON 3.0) on Smart Computing for Innovation and Advancement in Industry 4.0, scheduled to take place from 5-7 June 2024 at OP Jindal University, Raigarh-496109, Chhattisgarh, India (Hybrid Mode).

Please follow the Registrations Steps:

1. To know registration fee applicable, please visit to https://opjuieeeotc.in/ and go to Registration Menu. For above 6 pages (IEEE Template), INR 1000 / $100 per page will be applicable.

2. You can Pay Registration Amount Through Direct Payment Link as Follows (For Indian Author Only):
https://smartpay.easebuzz.in/125084/OTCON30

or

You can deposit/online/UPI to the following Bank Account (Indian and Foreign Authors):

Account Details:
A/C Name: OP Jindal University
A/c No.: 50200054124691
RTGS/NEFT IFSC: HDFC0001454
MICR: 496240002
Swift Code: HDFCINBBXXX

3. After Payment, Please Fill the following google form (Upload Cyber Receipt and Transaction Details):
https://forms.gle/BiRBQM8r8UF9MimE9

PAPER NAME

Yati_Final_thesis_chapter_document.doc x

WORD COUNT	CHARACTER COUNT
12780 Words	74556 Characters
PAGE COUNT	FILE SIZE
39 Pages	1.2MB
SUBMISSION DATE	REPORT DATE
May 12, 2024 8:56 PM GMT+5:30	May 12, 2024 8:57 PM GMT+5:30

12% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

- 6% Internet database
- 6% Publications database
- Crossref database
- Crossref Posted Content database
- 10% Submitted Works database

Excluded from Similarity Report

- Bibliographic material

How much of this submission has been generated by AI?

0%

of qualifying text in this submission has been determined to be generated by AI.

Caution: Percentage may not indicate academic misconduct. Review required.

It is essential to understand the limitations of AI detection before making decisions about a student's work. We encourage you to learn more about Turnitin's AI detection capabilities before using the tool.

Frequently Asked Questions

What does the percentage mean?
The percentage shown in the AI writing detection indicator and in the AI writing report is the amount of qualifying text within the submission that Turnitin's AI writing detection model determines was generated by AI.

Our testing has found that there is a higher incidence of false positives when the percentage is less than 20. In order to reduce the likelihood of misinterpretation, the AI indicator will display an asterisk for percentages less than 20 to call attention to the fact that the score is less reliable.

However, the final decision on whether any misconduct has occurred rests with the reviewer/instructor. They should use the percentage as a means to start a formative conversation with their student and/or use it to examine the submitted assignment in greater detail according to their school's policies.

How does Turnitin's indicator address false positives?
Our model only processes qualifying text in the form of long-form writing. Long-form writing means individual sentences contained in paragraphs that make up a longer piece of written work, such as an essay, a dissertation, or an article, etc. Qualifying text that has been determined to be AI-generated will be highlighted blue on the submission text.

Non-qualifying text, such as bullet points, annotated bibliographies, etc., will not be processed and can create disparity between the submission highlights and the percentage shown.

What does 'qualifying text' mean?
Sometimes false positives (incorrectly flagging human-written text as AI-generated), can include lists without a lot of structural variation, text that literally repeats itself, or text that has been paraphrased without developing new ideas. If our indicator shows a higher amount of AI writing in such text, we advise you to take that into consideration when looking at the percentage indicated.

In a longer document with a mix of authentic writing and AI generated text, it can be difficult to exactly determine where the AI writing begins and original writing ends, but our model should give you a reliable guide to start conversations with the submitting student.

REGISTRATION DETAILS

Yati Piplani, Pawan Singh Mehra, “Exploring Machine Learning Techniques for Enhanced Ayurvedic Doshas Diagnostics: A Comprehensive Review “The paper has been accepted in 4th International Conference on Machine Learning and Big Data Analytics (ICMLBDA2024) May 9-11, 2024, Indexed by Scopus. Paper Id: 144.

Receipt from I.A.A.S.S.E.

Receipt #1551-7141

AMOUNT PAID	DATE PAID	PAYMENT METHOD
₹8,000.00	Apr 6, 2024, 6:01:36 PM	VISA - 3633

SUMMARY

Full Time students (Non IAASSE Members) × 1	₹8,000.00
Amount charged	**₹8,000.00**

Paper Title: Exploring Machine Learning Techniques for Enhanced Ayurvedic Doshas Diagnostics: A Comprehensive Review
Paper ID: 144

Yati Piplani, Pawan Singh Mehra, “An Unsupervised Learning Model for Ayurvedic Prakriti Determination Using Smart Device Data” The paper has been accepted in at OPJU International Technology Conference (OTCON 3.0) on Smart Computing for Innovation and Advancement in Industry 4.0 June 5-7 2024, Indexed by Scopus in IEEE. Paper Id: 820.

Easebuzz

Paid Successfully To

₹7000.0
via UPI

OP JINDAL UNIVERSITY
ESTD. 2014
OPJU

https://www.opju.ac.in/

Your payment for Easebuzz ID E2404277SZBBDL successfully completed.

Hi Yati Piplani,
Thanks for using Easebuzz!
For any further queries you can get in touch with the merchant on rahul.thawait@opju.ac.in.

Transaction ID:
ECX8SW65HZ

Email :
yatipiplani07@gmail.com

Contact Number:
6388991376