Bridging Ayurveda and AI: Data Standardization for Improved Machine Learning Application

Dr.Sudha.P
Dept.of Computer Science Engineering
Presidency University
Bangalore,India
sudha.p@presidencyuniversity.in

Sanjana.B.K
Dept.of Computer Science Engineering
Presidency University
Bangalore,India
sanjanagowdabk@gmail.com

Thejaswini.G.K
Dept.of Computer Science Engineering
Presidency University
Bangalore,India
thejaswinigk23@gmail.com

Akash.K.A
Dept.of Computer Science Engineering
Presidency University
Bangalore,India
akashtej612 @gmail.com

***Abstract*—Ayurveda integration with machine learning (ML) applications must be grounded on a standardized and organized dataset to handle the complexity and heterogeneity of traditional medical terminology. The present paper suggests a process of data standardization in response to the vagueness in Ayurvedic texts to ensure uniformity in disease, symptom, treatment, and dosha categorization. A pre-defined ontology translated raw Ayurvedic terms into standardized terms to improve data quality for ML training. To analyze the impact of standardization, different ML models—Naïve Bayes, CNN, and BERT—were trained on standardized data. The results show that the maximum classification accuracy (100%) was achieved by BERT, which demonstrates the effectiveness of contextual embeddings for Ayurvedic text classification. The findings demonstrate that standardization significantly improves the performance of models, improving knowledge retrieval and compatibility with modern healthcare systems. This research contributes to building robust, machine-usable Ayurvedic datasets for AI-based diagnosis and treatment recommendation in traditional medicine.**

Keywords—Ayurveda, Machine Learning, Data Standardization, BERT, NLP, Knowledge Retrieval

I. INTRODUCTION

Ayurvedic medicine, having thousands of years of history, has its past extensively documented in literary sources. With the arrival of machine learning (ML) and artificial intelligence (AI), Ayurvedic knowledge is again put under the microscope to unlock meaningful insights through the process of digitization and analysis. Incorporation of Ayurvedic into modern ML systems faces colossal challenges driven mostly by text heterogeneity of structure, usage of variant terms, and non-existence of uniform data sets. Unstructured content of Ayurvedic literature restricts developing high-performance ML models, consequently hindering application of the model in clinical treatment and research work as well.

Despite attempts to automate Ayurvedic texts, the available datasets suffer from redundancy, vagueness, and the absence of a standardized vocabulary, which makes it difficult to build robust machine learning models. Further, differences in the description of medicinal plants, medicines, and diseases contribute to the nature of this problem. Hence, there is a pressing need to standardize Ayurvedic data in a systematic way to support machine learning-based applications.

This study proposes a data standardization framework employing Natural Language Processing (NLP) to enhance the quality and usability of Ayurvedic text-based data. Named Entity Recognition (NER), text classification, and ontology mapping are employed for data pre-processing and standardization of representation and vocabulary for uniformity. We also evaluate the accuracy of machine learning models trained on standardized data to see how different models perform. By resolving these challenges, this work contributes to the generation of high-quality, well-structured Ayurveda-pertinent datasets to be used in training machine learning algorithms. The quality of the output provided by this work could contribute notably to the integration of Ayurveda in AI-supported health systems, thereby contributing to predictive analytics, knowledge acquisition, and decision-making support in traditional medicine.

II. LITERATURE REVIEW

AI will be helpful for predictive modeling of Ayurvedic preparations, photochemical profiling, and efficacy optimization of herbal drugs. But for its full potential, concerns like the absence of standardized data sets, regulatory challenges, and compatibility with contemporary healthcare restrict it. These can be overcome through standardization of data, AI-based data banks, and interdisciplinary to augment scientific validation and worldwide acceptability of Ayurveda.[1]

This research examines the augmentation of Ayurvedic diagnosis through Multinomial Naïve Bayes (MNB) and K-modes clustering for examining Prakriti types and Dosha overlapping. The MNB classifier is used in predicting a person's Prakriti based on textual and categorical health features using probabilistic modeling. At the same time, K-modes clustering for categorical data detect underlying patterns in Dosha mixtures, discovering overlapping trends among Vata, Pitta, and Kapha.[2]

The article "Challenges and Opportunities in Natural Language Processing for Clinical Data" presents NLP's ability to derive significant insight from clinical unstructured text, like electronic health records (EHRs), to maximize healthcare outcomes. It also lays out some of the most compelling challenges, ranging from data privacy concerns to varied linguistic and contextual variations. In spite of such challenges, the paper also presents opportunities in the form of advances in deep learning, growing availability of structured health data, and improved interoperability standards, which can contribute towards augmenting clinical NLP applications. [3]

The application of machine learning (ML) in Ayurveda, the traditional health science of India, is suggested to modernize and enhance its diagnostic, predictive, and therapeutic functionalities. ML techniques can be applied to classify vast Ayurvedic databases and provide individualized Prakriti (body constitution) classification, disease prediction, and herbal treatment suggestion based on past patient records. Methods such as natural language processing (NLP) help to digitize ancient Ayurvedic manuscripts, and artificial intelligence-based models help in standardization and authentication of Ayurvedic formulations for evidence-based practice. With the combination of ancient wisdom and contemporary technology, ML can potentially make Ayurveda more accessible, precise, and acceptable globally within integrative healthcare systems.[4]

AI-driven models can improve Ayurvedic Prakriti (body constitution) analysis, streamline herbal formulations, and enable data standardization for clinical trials. Nonetheless, impediments like the unstructured nature of datasets, variance in conventional knowledge, low digitalization of Ayurvedic texts, and the necessity for regulatory frameworks deter the adoption of AI. Overcoming these gaps via data standardization, inter-disciplinary research collaboration, and AI-based research has the potential to unlock Ayurveda's ability to realize its applications in modern healthcare.[5]

The paper "Machine Learning Models Used for Prakriti Identification Using Prasna Pariksha in Ayurveda – A Review" discusses the application of machine learning (ML) methods to computerize the evaluation of Prakriti, a person's innate constitution in Ayurveda, based on Prasna Pariksha, a conventional questionnaire-based approach. The review identifies different ML algorithms such as Support Vector Machines (SVM), K-Nearest Neighbors (KNN), Naive Bayes (NB) classifiers, Artificial Neural Networks (ANN), Decision Trees, and ensemble approaches such as AdaBoost used for Prakriti type prediction. Importantly, research has attained high accuracy levels, for example, 97% utilizing AdaBoost with hyper parameter optimization, which establishes the potential of ML in the augmentation of personalized medicine in the Ayurvedic system.[6]

Analysis of Big Data in conventional knowledge-based Ayurveda medicine has meant utilizing cutting-edge computational methods to analyze vast pools of historical and modern Ayurvedic data. Using machine learning, natural language processing, and statistics, scientists can identify implicit patterns, confirm ancient intuition with contemporary scientific proof, and improve on personalized treatment strategies. This method enables the integration of Ayurveda with contemporary healthcare through standardization of datasets, enhanced diagnosis and prognosis models, and maximized treatment effectiveness. It also closes the gap between ancient Ayurvedic literature and evidence-based medicine, promoting international acceptance and innovation in holistic healthcare.[7]

Natural Language Processing (NLP) is transforming Ayurveda by making it possible to extract, analyze, and interpret large volumes of ancient Sanskrit texts, clinical data, and patient information. Through the use of AI-powered NLP models, scientists can normalize Ayurvedic vocabularies, improve prakriti (body constitution) evaluation, and develop smart diagnostic systems. This combination of ancient knowledge with advanced AI makes personalized treatment suggestions, enhances predictive analysis, and improves compatibility with contemporary healthcare systems. With the ongoing development of NLP, it can potentially close the gap between Ayurveda and evidence-based medicine and make holistic healing more accessible and data-based.[8]

A Semantic Annotation and Querying Framework for semi-structured Ayurvedic texts seeks to widen the accessibility and understandability of ancient medical knowledge by applying Natural Language Processing (NLP) and ontology-based techniques. The framework entails annotating Ayurvedic texts with metadata in structured form, associating concepts like Prakriti (body constitution), Doshas, herbs, and treatments with a clearly defined ontology. By combining semantic search and query mechanisms, the researchers and practitioners are able to fetch accurate and context-sensitive information, transcending limitations presented by unstructured textual data. Such a mechanism enhances interoperability to a greater extent, assisting clinical decision support, personalized medicine, and AI-facilitated analysis of Ayurvedic literature.[9]

This article centers on the importance of AI for enhancing pharmaceutical formulation, quality assurance, and standardization in the Ayurvedic healthcare sector. With increased demand for Ayurvedic medicine globally, AI provides means of ensuring efficacy, ensuring uniformity, and accelerating drug discovery. AI can utilize automation and predictive analysis to streamline the complex processes involved in Ayurvedic pharmaceutics. This can help solve issues of raw material variance, dose consistency, and shelf life. The research illustrates how AI can facilitate personalized therapy, enhance regulatory compliance, and facilitate the integration of Ayurveda into modern medical practices.[10]

This research introduces the use of pharmacogenomics and ayurgenomics in personalized medicine. When put into perspective using Ayurvedic concepts, it shows how genetic diversity can enhance the quality of care. The research showcases how Ayurvedic distinctions such as Prakriti and genetic information can be employed to individualize care. The research imagines a utopian world society wherein individualized therapy blends the finer points of Ayurveda with the latest genetics to ensure more effective and safe therapy. [11]

To be able to deliver individualized treatment, this research discusses the possible convergence of genetics with Ayurvedic concepts, e.g., Prakriti (constitution of the individual). This discusses how genomic science can inform Ayurvedic classifications and perhaps establish scientifically validated individualized treatment regimens. The study introduces a model for integrating Ayurveda and contemporary medical research by exploring genetic markers in relation to Prakriti. This juncture has the purpose of progressing Ayurveda as an ever-more popularly accepted medical tradition from a science perspective within the global health delivery system via innovative channels of individually tailored treatment options, dietary advice, and preventive medicine systems.[12]

This article discusses the incorporation of Artificial Intelligence (AI) in Ayurveda, a 5,000-year-old traditional medical system. Ayurveda focuses on a holistic and personalized system of health that is concerned with the balance of body, mind, and spirit. The research discusses how AI can be used to improve Ayurvedic diagnosis, treatment, and personalized medicine through sophisticated data analytics, machine learning, and predictive modeling. A review of literature in the form of research papers, ancient Ayurvedic treatises, and specialist views was undertaken to assess the role of AI in Ayurveda. AI can be used to scan huge volumes of past and current data to discern health trends, foretell vulnerability to diseases in accordance with Prakriti (body constitution), and enhance treatment strategies. Moreover, AI-based systems can help to unravel novel herbal recipes, thus augmenting Ayurvedic therapy possibilities. The integration of AI and Ayurveda has the potential to transform healthcare through the blending of ancient knowledge and contemporary technology, resulting in more efficient, tailored, and preventative medical interventions.[13]

To ensure the safety, efficacy, and global acceptability of Ayurvedic products, the present study underlines the absolute necessity of standardization. It brings into focus issues that concern the authenticity of Ayurvedic drugs, e.g., variability in raw material, lack of standards, and adulteration. Enhancements in pharmacopoeial standards and quality control measures are proposed as potential solutions, as are regulatory systems like those of the Ministry of AYUSH and the WHO. The research also highlights how important it is to bridge traditional wisdom with modern scientific evidence to make Ayurveda a dignified global healthcare system [14]

This paper discusses the current scenario of Ayurvedic research and the main hurdles for scientific validation. The paper refers to problems like unsystematic formulation, absence of clinical trials, and the challenge of integrating scientific concepts with Ayurvedic principles. The publication deals with upgraded clinical study paradigms, application of Good Manufacturing Practices (GMP), and computerization of Ayurveda manuscripts. It also highlights the importance of interdisciplinary research collaboration for strengthening knowledge and worldwide recognition of Ayurveda. The project promises a systematic approach of formulation and authentication of Ayurvedic medicine through integration of traditional knowledge with modern research ethics.[15]

III. Proposed approach

The data extraction, preprocessing, and standardization of Ayurveda medical terminologies for making them machine-readable in ML applications were carried out through a concatenated approach. The key stages are:

A. *Data Collection & Preprocessing:*

The dataset utilized in this research was downloaded from Kaggle (Ayurvedic Formulations and Their Indications), which contains in-depth information regarding various Ayurvedic formulations, the ingredients they are comprised of, and the specific diseases they are prescribed for. This dataset is a great source in understanding traditional Ayurvedic medicine and its application in real-life situations. To verify whether the dataset is appropriate for Natural Language Processing (NLP) tasks, a series of preprocessing methods were applied. Initially, tokenization was employed to divide the text into significant words. Subsequently, stopword removal was applied to eliminate frequent words that do not improve the understanding of Ayurvedic entities. Furthermore, lemmatization was applied to transform words to their base forms so that the NER and classification tasks would be uniform. These preprocessing methods rendered the text data more organized and susceptible to additional analysis in text classification and ontology mapping.

B. *NLP Techniques:*

1) *Named Entity Recognition (NER):*
NER was employed to discover crucial Ayurvedic entities such as medicinal plants, formulations, diseases, and therapeutic properties from free text. A highly filtered corpus of Ayurvedic dataset was trained on standard terminologies with entity tags marked. SpaCy and transformer models were used for the identification of entities. NER output was used to identify and tag relevant Ayurvedic terms as the starting point for text standardization in the next step.

2) *Text Categorization:*
We used a multi-label text classification model to label complete Ayurvedic texts into pre-defined categories: DISEASE, DOSHA, TREATMENT, and HERB. SpaCy's textcat_multilabel pipeline was utilized to train the classification model. A small dataset where each sample of text was labeled with a single or group of category tags was utilized for training the model. Training was done using an iterative process of 10 epochs to train the model's performance. The model trained was tested on a test sample sentence, "Ashwagandha aids in stress relief," which gave the following predictions:
DISEASE - 0.9216
DOSHA - 0.1383
TREATMENT - 0.0835
HERB - 0.5697

3) *Ontology mapping :*
Was used to standardize and categorize the extracted terms by mapping them against pre-defined Ayurvedic categories. The process started with defining a structured ontology that grouped the terms into DISEASE, DOSHA, TREATMENT, and HERB, maintaining classification consistency. For example, terms such as "Netraroga" were mapped into DISEASE, whereas "Vatavikara" fell under DOSHA. For this, TF-IDF vectorization was employed to transform symptom descriptions from the dataset into numerical form. The cosine similarity was then computed between the symptom vectors and pre-defined ontology terms to identify the best-matching

category. If the similarity measure is greater than a predefined threshold value (say, 0.3), the symptom was categorized into the respective category; otherwise, it was categorized under UNKNOWN. This strategy guaranteed that Ayurvedic symptoms and concepts were mapped in a systematic manner, improving data standardization for subsequent machine learning use.

TABLE 1 : Ontology Mapping

Symptom	**Description**	**Mapped Category**
Vatavikara	Disorders related to the Vata dosha	Dosha
Vrana	Wound/injuries	Disease
Netraroga	Eye disorders	Disease
Galaganda	Goiter	Unknown

C. Dataset Standardization Approach:

1) Entity Categorization:

- Symptoms – Symptoms may be classified into broad Ayurvedic categories according to dosha imbalances (Vata, Pitta, Kapha).
- Diseases – Diseases of Ayurveda are often defined in Sanskrit. Standardization would require relating them with standardized modern medical or Ayurvedic terminologies for easier categorization.
- Treatments – Ayurvedic treatments typically involve herbal formulations, dietetic changes, and therapies (like Panchakarma).
- Dosha – As Ayurvedic treatment and diagnosis rely significantly on dosha constitution, dosha-based symptom and treatment definition and mapping must be standardized.

Before the application of the dataset standardization method, the Ayurveda terminology contained orthography, classification, and explanation inconsistencies that resulted in inconsistencies in machine learning procedures. The raw dataset contained 300 entries with many uncertain or unclassified terms that were classified as UNKNOWN. Following standardization, vague terms were matched against determined categories built in accordance with Ayurvedic ontology, lowering the percentage of UNKNOWN classifications.

D. ML Model Implementation:

1) CNN Model:

To classify the Ayurvedic symptoms and normalize the dataset, a Convolution Neural Network (CNN) model was employed. CNNs are especially suitable to learn spatial hierarchies from text data when given in sequential forms and therefore are highly effective for text categorization-based tasks.

a) Dataset Preparation:

The dataset of Ayurvedic text was tokenized for the first time with the Keras Tokenizer class, thus transforming every word into numeric value. A vocabulary of 5000 words was used to incorporate the most descriptive vocabulary, and the output sequences of tokens were padded to a common maximum length of 50 words to supply a uniform input size. The dataset was split into training and test sets with 80% being used for training and 20% being used for testing.

b) Architecure of the CNN model are:

- The embedding layer converts tokenized lexical items into dense vector representations.
- 1D Convolution Layer: Uses 128 filters with a size 5 to spot patterns in series of words.
- Global Max Pooling Layer: It reduces dimensionality without sacrificing the most important features.
- Fully Connected Dense Layer: 64 neurons with ReLU activation.
- The output layer uses a SoftMax activation function to classify text within each Ayurvedic category.

c) Training Process:

The model was developed utilizing categorical cross-entropy as the target loss function along with the Adam optimizer to allow easy gradient adjustments. The data was split into training and validation subsets in an 80-20 ratio. Model training was conducted for 10 epochs with a batch size of 16.

2) BERT Framework:

Pre-trained Model Fine-Tuning Modern NLP model BERT (Bidirectional Encoder Representations from Transformers) is designed to learn context bidirectional.

a) Data Preparation:

Dataset Preparation: Tokenized inputs with attention masks are needed to process padding efficiently, and these are what BERT needs. Ayurveda symptoms and labels were tokenized through the Bert-base-uncased tokenizer.

b) Model Fine-Tuning:

A pre-trained Bert-base-uncased was fine-tuned specifically for classifying Ayurvedic texts. The dataset was divided into the training and validation subsets, and the model was trained for several epochs with an optimizer of the transformer architecture.

c) The tuning process involved:

- Objective Function: Sparse Categorical Cross-Entropy
- Optimizer: AdamW
- Batch Size: 8
- Epochs: 5

3) Naïve Bayes:

Naïve Bayes is a probability-based classifier. It is a highly efficient and simple technique for text categorization as it treats a document as word-wise independent.

a) Dataset Preparation:

The dataset was term-frequency-inverse document frequency (TF-IDF) vectorized in order to convert the text to numerical features. Labels were changed into their corresponding classes for classification.

b) Model Training:

A Multinomial Naïve Bayes classifier was used to train the vectorized dataset. Accuracy and classification reports were employed to validate the model with the actual and predicted labels. Prediction from novel Ayurvedic text samples identified misclassification, hence supporting its drawback

IV. RESULT AND DISCUSSION

A. *Standardized Dataset Presentation*

The standardized dataset has a better structure and mapping of Ayurvedic medical terminologies, which provides uniform representation from different sources. The standardization process entailed the alignment of synonyms, fixing spelling differences, and mapping each entity to a pre-defined ontology for uniformity in ML applications. After the standardization the number of "UNKNOWN" were decreased from 287 to 0.

This change guarantees that such variations like "Shotha," "Sotha," and "Inflammation" are placed under one label, avoiding redundancy and ambiguity. Standardization also enables ICD-equivalent mapping, making Ayurvedic terms more compatible in today's medical databases.

B. *ML Model Performance Comparison*

To evaluate the effectiveness of different machine learning models for Ayurvedic text classification, we compared CNN, BERT, and Naïve Bayes using key performance metrics: Accuracy, Precision, Recall, and F1-score.

TABLE 2: Performance Metrics Table

Model	Accuracy	Precision	Recall	F1-score
BERT	1.00	1.00	1.00	1.00
CNN	0.80	0.75	0.75	0.75
Naïve Bayes	0.50	0.25	0.50	0.33

BERT outperformed both CNN and Naïve Bayes, achieving 100% accuracy with perfect classification, making it the most suitable model for Ayurvedic text classification. CNN showed promising results with 80% accuracy, but its performance can be further improved with additional training data and hyper parameter tuning. In contrast, Naïve Bayes struggled significantly, achieving only 50% accuracy due to its reliance on word frequency, which led to frequent misclassification of Ayurvedic terms. Deep learning models (BERT & CNN) handled domain-specific text much better than traditional models like Naïve Bayes, making them more reliable for standardizing Ayurvedic knowledge.

Standardization improved accuracy and speed of Ayurvedic text classification by incorporating deep learning approaches, like BERT and CNN. Such deep learning methods were proved to enhance classification accuracy as contextual meanings of words can be correctly extracted, with lower chances of misclassification. On the other hand, conventional models such as Naïve Bayes had a problem with variations in Ayurvedic texts, underscoring the significance of uniform data for training effective models. The formal dataset further facilitated efficient retrieval of knowledge so that symptom, disease, and treatment mapping could be accurately achieved. In total, the integration of data standardization and sophisticated ML methods fortifies the basis for AI-based Ayurveda applications to provide more trustworthy and consistent results in healthcare and research.

V. RESULT AND DISCUSSION

This research proved that data standardization enhances Ayurvedic text classification for machine learning use. Standardization of critical entities like treatments, Doshas, and diseases resulted in improved model performance with deep learning models like CNN and BERT over conventional methods such as Naïve Bayes. The standardized dataset facilitated accurate mapping, enhancing knowledge retrieval and AI-based analysis in Ayurveda. However, some limitations were noted, such as a fairly small dataset size and possible omissions in the standardization of rare or compound Ayurvedic words. The performance of the model was mixed depending on the complexity of input, suggesting that better contextual awareness is needed. Future research will involve increasing dataset coverage through the inclusion of more Ayurvedic texts and research articles, improving ontology integration to enhance knowledge representation, and investigating more sophisticated NLP methods such as transformer-based models with domain-specific fine-tuning. Such advancements will lead to more robust AI-based solutions in Ayurvedic healthcare, benefiting both practitioners and researchers.

REFERENCES

[1] A. V. Arun, P. N. Balasaheb, J. V. Babasaheb, J. D. Kailas, K. R. Adinath, and N. D. Dadasaheb, "Artificial intelligence and challenges in Ayurveda pharmaceutics: A review," Research Journal of Science and Technology, vol. 16, no. 3, pp. 237–244, 2024.

[2] P. Bidve and S. Mishra, "Enhancing Ayurvedic Diagnosis using Multinomial Naive Bayes and K-modes Clustering: An Investigation into Prakriti Types and Dosha Overlapping," arXiv preprint, arXiv:2310.02920, 2023.

[3] G. Névéol, K. B. Zweigenbaum, S. R. Velupillai, W. Chapman, M. Suominen, and P. Savova, "Challenges and Opportunities in Natural Language Processing for Clinical Data," Journal of Biomedical Semantics, vol. 9, no. 1, p. 1, 2018.

[4] Jadhav Vikas, S., Wakale Ashwini, D. and Mane, S.R., Integration of machine learning in Ayurveda: An Indian traditional health science.

[5] T. M. Nesari, "Artificial intelligence in the sector of Ayurveda: Scope and opportunities," Int. J. Ayurveda Res., vol. 4, no. 2, pp. 57–60, 2023.

[6] L. Bheemavarapu and K. U. Rani, "Machine learning models used for Prakriti identification using Prasna Pariksha in Ayurveda–A review," Mathematical Statistician and Engineering Applications, vol. 72, no. 1, pp. 1942–1951, 2023.

[7] H. Singh, S. Bhargava, S. Ganeshan, R. Kaur, T. Sethi, M. Sharma, M. Chauhan, N. Chauhan, R. Chauhan, P. Chauhan, and S. K. Brahmachari, "Big data analysis of traditional knowledge-based Ayurveda medicine," Progress in Preventive Medicine, vol. 3, no. 5, p. e0020, 2018.

[8] A. M. S., "Ayurveda Meets AI: How NLP Is Shaping the Future of Holistic Medicine," Journal of Emerging Technologies and Innovative Research (JETIR), vol. 11, no. 6, p. JETIR2406A62, Jun. 2024.

[9] H. Terdalkar, A. Bhattacharya, M. Dubey, and B. N. Singh, "Semantic Annotation and Querying Framework based on Semi-structured Ayurvedic Text," arXiv preprint, arXiv:2202.00216, 2022

[10] A. V. Arun, P. N. Balasaheb, J. V. Babasaheb, K. J. D. Kailas, A. K. R. Adinath, and D. N. Dadasaheb, "Artificial Intelligence and Challenges in Ayurveda Pharmaceutics: A Review," Research Journal of Science and Technology, vol. 16, no. 3, pp. 237-244, 2024.

[11] P. D. Gupta, “Pharmacogenetics, pharmacogenomics and ayurgenomics for personalized medicine: A paradigm shift,” Indian Journal of Pharmaceutical Sciences, vol. 77, no. 2, pp. 135–141, Mar.–Apr. 2015, doi: 10.4103/0250-474X.156543.

[12] Y. D. Madgulwar and K. J. Shewalkar, “The intersection of Ayurveda and genomics: Exploring Ayurgenomics for personalized health solutions,” Journal of Ayurveda and Integrated Medical Sciences, vol. 9, no. 10, pp. 168–177, 2024.

[13] Sanjay Gupta, Narasimha V, Vijaya Lakshmi A. Artificial Intelligence (AI) in Ayurveda: Its Application and Relevance. Ayushdhara [Internet]. 2025Jan.15 [cited 2025Mar.10];11(6):165-9.

[14] Jyothi Raga P M, Vivek P, & Harinarayanan C M. (2023). Need of Standardization of Ayurveda Formulations. International Journal of Ayurveda and Pharma Research, 11(10), 40-45.

[15] A. Chauhan, D. K. Semwal, S. P. Mishra, and R. B. Semwal, “Ayurvedic research and methodology: Present status and future strategies,” AYU (An International Quarterly Journal of Research in Ayurveda), vol. 36, no. 4, pp. 364–369, 2015.

[16] Raghav Singh, “Ayurvedic Formulations and their Indications,” Kaggle,2023. https://www.kaggle.com/datasets/raghavdecoded/ayurvedic-formulations-and-their-indications