



RAG-LLM Powered Multi-Dataset Machine Learning Framework for Personalized Ayurvedic Migraine Management

Sakshi Hedke¹, Saniya Samant¹ and Laxmi Bewoor² *

¹ Vishwakarma Institute of Information Technology, Computer Engineering Department, Pune, India

² Vishwakarma Institute of Technology, Computer Engineering Department, Pune, India

sakshi.22210064@viit.ac.in, saniya.22210224@viit.ac.in,
laxmi.bewoor@vit.edu*

Abstract. Past several decades millions of people are suffering from various neurological disorders and migraine is one of the severe condition affecting overall population. Normally the medication includes allopathic drugs which provide temporary relief but unable to handle root cause and use of such drugs for longer time is life threatening. Ayurveda, historical roots from ancient Indian system has been proven to be better alternative medicine system in chronicle diseases. Tridosha analysis is the foundation behind Ayurvedic treatment which consists of determining imbalances in Dosha viz. Vata, Pitta and Kapha, Based on the Prakriti of human being personalized treatments are recommended which not only look into symptoms but also address the root causes of the problem. However, this treasure is not yet conceptualized in an appropriate intelligent digital health system. This work proposes a hybrid framework integrating Machine Learning (ML), Retrieval-Augmented Generation (RAG), and Large Language Models (LLMs) which provides personalized Ayurvedic recommendations to deal with migraine. Ensemble classifiers predicts Dosha imbalance based on patient symptoms and lifestyle features whereas RAG is extremely useful for retrieving relevant Ayurvedic therapies, herbal formulations and dietary guidance. Groq-powered LLaMA3 model uses this knowledge is for providing human readable recommendations. The proposed work has synthetic dataset of 5000 patients so it should be thought of as clinical diagnosis supporting tool not a substitute of diagnosis. It is an attempt to provide a research prototype which is perfect blend of traditional knowledge and contemporary technique demonstrating personalized, interpretable healthcare decisions.

Keywords: Migraine, Ayurveda, Retrieval Augmented Generation, Large Language Models, Machine Learning, Dosha Prediction.

1 Introduction

Migraine is a neurological disorder affecting approximately 14% of global population characterized by recurrent, moderate-to-severe throbbing headaches often accompanied by nausea, vomiting, and sensitivity to light, sound, and smells. Normally clinical practitioners prescribe allopathic drugs viz. NSAIDs, triptans, prophylactics with incomplete relief and an increasing burden of side effects. Allopathic recommendations are reactive and symptom focused, and often failed to interrogate the deeper physiological or lifestyle imbalances that sustain the condition. Ayurveda, an

indigenous ancient medical science of India, is science of health based on universal principles and deep insights into the linking between mind and body and the laws of nature which structure all progress in life and more specially deals with the root cause of any disease. It has been devised that Ayurvedic diagnosis is knowledge-intensive majorly focusing on tridoshas viz. Vata, Pitta and Kapha, highly personalised, and difficult to scale without the aid of intelligent computation. This motivated us to develop the system which should integrate the ancient knowledge leveraging the contemporary technique.

As per Ayurveda, migraine resembles to “Ardhavabhedaka”, causes due to imbalance between the three basic biological forces (Vata, Pitta and Kapha) indicating unique symptom based on individual predominance of Dosha and hence require specific therapeutic strategies. Person with Vata-predominant migraine usually suffers from throbbing, anxiety, and sleep disturbance. The symptoms for Pitta-type individuals are burning intensity, photophobia, and sensitivity to heat. Kapha-type individuals suffer from dull heaviness, congestion, and sluggishness. The clinical benefits like reduction in headache frequency, better sleep quality, reduced stress are witnessed with the reviewed research. Recommended Ayurvedic therapies such as Nasya Karma (medicated nasal administration), Shirodhara (rhythmic oil-streaming on the forehead), Panchakarma detoxification, and specific herbal formulations such as Pathyadi Kwatha and Laghu Sutashekhar Rasa [1, 2, 3] are proven to be the best treatments with long relief and without any side effect.

In spite of this, Ayurvedic knowledge is not well integrated into modern digital health tools. Domain curated data and well-structured Ayurvedic dataset is the major challenges in this research. So development of such a curated data is the primary focus of this work. Along with that current progress in AI, particularly in Large Language Models (LLMs) and Retrieval-Augmented Generation (RAG), offer a promising path to extend this research. As LLMs are susceptible to hallucination, it might generate some irrelevant medical text as they are not trained with domain knowledge, which is a significant liability in clinical settings. However RAG addresses this issue by grounding generative outputs in domain-specific knowledge bases retrieved in real time, leading to a significant improvement in factual accuracy and interpretability [4, 5]. So, in this work, we propose a hybrid AI framework that combines (1) ML-based Dosha classification from symptom and lifestyle inputs, (2) multi-dataset Ayurvedic knowledge integration, (3) FAISS-based semantic retrieval with Sentence Transformer embedding, and (4) Groq-powered LLaMA3 for personalized recommendation generation. The system is designed to be interpretable, scalable and rooted in classical Ayurvedic principles.

The main contributions of this research are:

- (1) Development and compilation of five datasets in the domain of ayurveda for dosha prediction, therapy mapping and knowledge retrieval.
- (2) Develop and benchmark ML models (Logistic Regression, SVM, Random Forest, XGBoost) for dosha classification.

- (3) Full RAG pipeline embedding, indexing, retrieval and generation with grounding to Ayurvedic knowledge.
- (4) A rule-based contraindication filtering layer that imposes ayurvedic safety constraints on the generated recommendations.

2 Literature Review

Both traditional and biomedical medicine have a great deal of research on the management of migraine. The current standard of care includes pharmacological approaches such as NSAIDs, triptans and preventive agents, but these have high rates of medication overuse, gastrointestinal adverse effects and limited efficacy for chronic presentations [6]. These shortcomings have perpetuated the interest in complementary approaches, of which Ayurveda is the most systematically documented in the Indian medical tradition.

Classical Ayurvedic texts such as the Charaka Samhita and the Sushruta Samhita describe Ardhavabhedaka as a dosha-mediated cranial disorder that can be treated with individualized therapeutic interventions. Nasya Karma not only decreases the intensity of headache but also reduces the associated symptoms such as photophobia, stress etc [3, 7]. Shirodhara has significant benefits in the modulation of sympathetic nervous system and regulation of sleep quality [8]. Herbal agents such as ginger, turmeric, Pathyadi Kwatha, Laghu Sutashekhara Rasa [9,10] proven to be the best natural healers of migraine because of its anti-inflammatory and neuroprotective properties. Panchakarma is identified and reported[11] as best detoxification method for sustained symptomatic improvement with very few adverse effects as compared to long-term pharmacotherapy.

The proposed work surveyed the literature from computational perspective too. It has been observed that standard machine learning based classifiers with complex feature interactions [12, 13] provides excellent performance for structured clinical datasets. Advanced Transformer based NLP models such as BERT [14] and the LLaMA capable of text understanding and generation in the medical domain. However, lack of reliable domain understanding lead to hallucination so relying on standalone LLMs won't provide the required quality of solution. Hence, building a efficient RAG frameworks spurred significant interest specially in healthcare [4, 5, 15, 16]. Few attempts like AyuRAG, and hybrid Ayurvedic recommendation systems contributed to develop Ayurvedic knowledge bases with embedding models and vector databases [17]. However, these systems face the limitation of fragmented datasets, no contraindication-aware safety layers and little interaction between prediction and explanation parts.

The reported research in Ayurvedic practise is not well documented and that led to a major limitation of small sample sizes, which in turn affects the lack of standardized datasets, and hence limited scalability. The proposed work is an attempt to fills this gap in the field. Additionally, it has been witnessed that translating this clinical experimentation into decision support system requires adequate computational infrastructure which is rarely thought of in this context. So, overall the proposed work

is a small attempt to address these gaps through unified multi-dataset architecture, integrated ML-RAG pipeline and a rule-based recommendation filtering mechanism.

3 Proposed System Architecture

The proposed framework provides a modular pipeline based architecture comprising of the symptoms collection through user input, Data Processing with knowledge integration, Prediction of Dosha using ML classifiers, Generation of semantic embedding, Knowledge Retrieval through similarity search, providing recommendations through LLM and personalized Ayurvedic treatment suggestion. The workflow is illustrated in Fig.1.

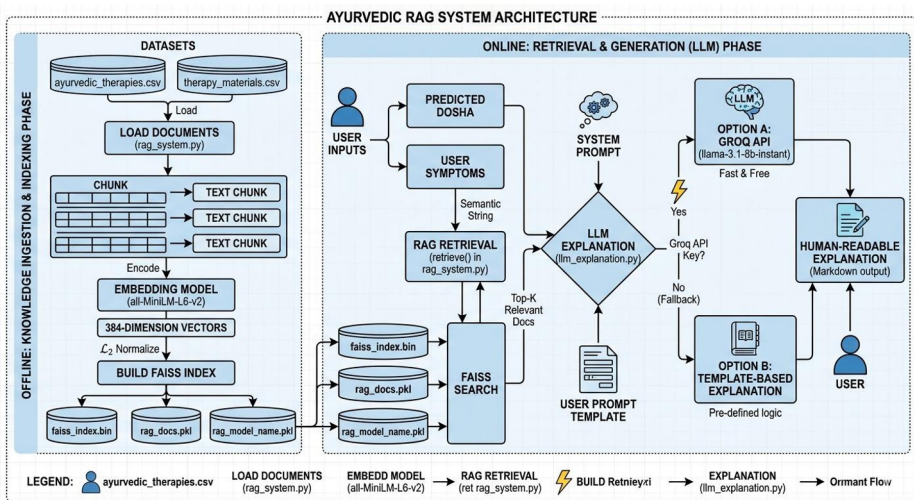


Fig. 1. Ayurvedic RAG based system Architecture

The workflow starts with capturing structured user input such as migraine-related symptoms, lifestyle habits, environmental triggers, sleep patterns, dietary behaviours, and headache characteristics. These features are pre-processed and fed into the trained ML classifiers for Dosha prediction. Based on the predicted Dosha, the knowledgebase is created for providing the Ayurvedic therapies. In order to do the same, system uses Sentence Transformer embedding to encode records of Ayurvedic knowledge and performs semantic similarity retrieval using an FAISS index. The retrieved therapy and herbal records along with the user symptoms and predicted dosha label are combined into a structured prompt that is passed to the Groq-powered LLaMA3 model to generate personalized recommendations. A rule based filtering layer is used to perform post-retrieval safety checks by filtering out contraindicated therapies based on dosha, symptom severity and seasonal context before final output rendering.

4 Experimental Design and Methodology

4.1 Dataset

One of the main hurdles in researching Ayurvedic healthcare is the almost complete absence of publicly available, structured datasets. To bridge this gap, we created five domain-specific datasets for this work. All these datasets alongwith its description is provided in this section.

Synthetic Patient Dataset. A synthetic dataset of approximately 5,000 patient records was generated to simulate realistic migraine presentations. Dataset was generated using a rule based generation method based on the classical ayurvedic diagnostic principles as described in the Charaka Samhita and Sushruta Samhita and peer reviewed clinical studies on Dosha-symptom relations.

Probabilistic feature assignment rules were defined for each dosha category (Vata, Pitta, Kapha) based on characteristic migraine symptoms, lifestyle factors and environmental triggers from published Ayurvedic and biomedical literature. The features examined included age, gender, headache duration, pain intensity, pain type, nausea, photophobia, sleep quality, stress level, screen exposure, seasonal triggers, dietary habits and lifestyle patterns. The dominant dosha category was the target class label.

Added controlled noise provide random variation of feature values, partial symptom overlap between Dosha boundaries and targeted label perturbation on a subset of samples to improve realism and reduce structural over-fitting. However, generated patterns were cross checked with this diagnostic framework to retain the internal consistency. Table 1 describes mapping of Doshas to its defining symptom profile along with key literature sources which are useful for feature generation.

Table 1. Dosha-symptom mapping.

Dosha	Characteristic Symptoms & Triggers	Key Literature Sources
Vata	Throbbing pain, anxiety, dry skin, sleep disturbance, cold sensitivity, constipation	[1, 3, 7, 10]
Pitta	Burning/sharp pain, photophobia, heat sensitivity, irritability, nausea, acid reflux	[2, 8, 9, 11]
Kapha	Dull heavy pain, congestion, fatigue, oversleeping, poor appetite, mucus accumulation	[3, 7, 11]

It is well understood that this instead of synthetic data generation real time data reviewed by formal clinical expert need to be curated. The main aim of proposed work is to provide a framework so that expert in this field can think of such design and

provide feedbacks regarding the same to build further collaborations. Such collaboration will be very useful for dataset refinement, external validation as well as spreading indigenous alternatives.

Therapy-Material Dataset. This dataset links Ayurvedic therapies to related materials, dosha targets, therapeutic effects and contraindications. The main fields are: Therapy name, Material used, Material type, Dosha targeted, Benefits, Contraindications. Therapies include Nasya Karma, Shirodhara, Basti, Dhumapana, and Panchakarma procedures. The rule-based recommendation filtering layer utilizes directly the contraindication fields.

Dosha-Symptom Mapping Dataset. This dataset describes logical symptom-to-dosha relations that are used to improve recommendation interpretability and to support retrieval filtering by predicted dosha class.

Ayurvedic Medicine Dataset. The data set comprises structured records of herbal formulations with active ingredients, dosha applicability, therapeutic effect, dosage guidance and usage recommendations.

Allopathy Comparison Dataset. In order to provide comparative analysis, dataset for allopathy based drugs prescribed for migraine and their side effects was maintained.

4.2 Methodology

Data Pre-processing

Pre-processing step is performed initially for all datasets. All categorical variables were encoded with label encoding. Imputation techniques were used to handle the missing values. MinMax scalar was used for feature normalization. Correlation and importance analysis was used for selecting the appropriate features. 80/20 train-test split was kept for model training. Feature ordering and encoding schemes kept consistent for training and inference stages.

Machine Learning-Based Dosha Prediction

Three ML classifiers were implemented and comparatively evaluated for three classes dosha prediction. We used Logistic Regression as the linear baseline. Random Forest was included for its robustness in ensemble, resistance to noise and capability to capture complex interactions between features. XGBoost was used for high performance gradient boosting classification. For all models, a 5-fold cross-validation strategy was applied for a reliable estimation of generalization.

Retrieval-Augmented Generation (RAG) Framework

Here is how the RAG pipeline was built. Structured Ayurvedic knowledge records from the therapy dosha-symptom, and medicine datasets were translated into natural language text representations. The texts were encoded with the Sentence Transformer model all-MiniLM-L6-v2 that creates 384-dimensional dense vector embedding capturing semantic relationships beyond keyword matching.

To perform an exact nearest-neighbour search with high retrieval precision, we built an index of all encoded records with the FAISS IndexFlatL2 index. At inference time, the

user symptom descriptions and predicted dosha labels were concatenated into a query, embedded using the same Sentence Transformer model, and matched against the index using cosine similarity. In response to each query, the top-5 most semantically relevant records (top-k = 5) were dynamically retrieved. The retrieved records, the symptom profile of the user, and the dosha label prediction were assembled into a structured prompt template and submitted to the Groq-powered LLaMA3 model. The prompt template is formatted as follows:

You are an Ayurvedic health advisor. The patient has been diagnosed with a [DOSHA] imbalance. Their reported symptoms include: [SYMPTOM LIST]. Based on the following retrieved Ayurvedic knowledge: [RETRIEVED CONTEXT], provide a personalized recommendation including: (1) Dosha explanation, (2) Recommended therapies, (3) Herbal remedies, (4) Dietary guidance, (5) Lifestyle modifications (6) Preventive suggestions. Ensure all recommendations are grounded in the provided Ayurvedic context.

Hallucination control is done through two mechanisms: (1) retrieval grounding bounds LLM generation to domain-validated Ayurvedic content; and (2) the rule-based filtering layer performs post-generation contraindication validation. The authors manually checked the generated recommendations by consulting classical Ayurvedic texts and published clinical literature for factual consistency. However as this is an attempt of building the prototype the next action is to validate the model with Ayurvedic Practitioner and then work with real data.

Rule-Based Recommendation Layer

A rule-based filtering mechanism was introduced after LLM generation for compliance with Ayurvedic safety. The engine uses contraindication fields in the Therapy-Material Dataset to exclude therapies that are contraindicated for the predicted Dosha, contraindicated based on the reported symptom severity, or seasonally inappropriate. This layer retains the classical Ayurvedic diagnostic logic and ensures that the system does not output potentially harmful recommendations. This layer can be considered as a clinically aware post-processing safety gate.

5 Results and Discussion

We tested the framework on four aspects i.e., Dosha prediction, retrieval, recommendation, and system interpretability. The proposed framework has been tested with four aspects specifically i) prediction performance by means of classification accuracy and related metrics, ii) retrieval effectiveness with relevant information identification and iii) recommendation quality.

5.1 Classifier Performance

Logistic Regression, Random Forest and XGBoost classifiers were used to predict the dosha for patient dataset. Fig. 2 provides the comparative analysis of all the classifiers.

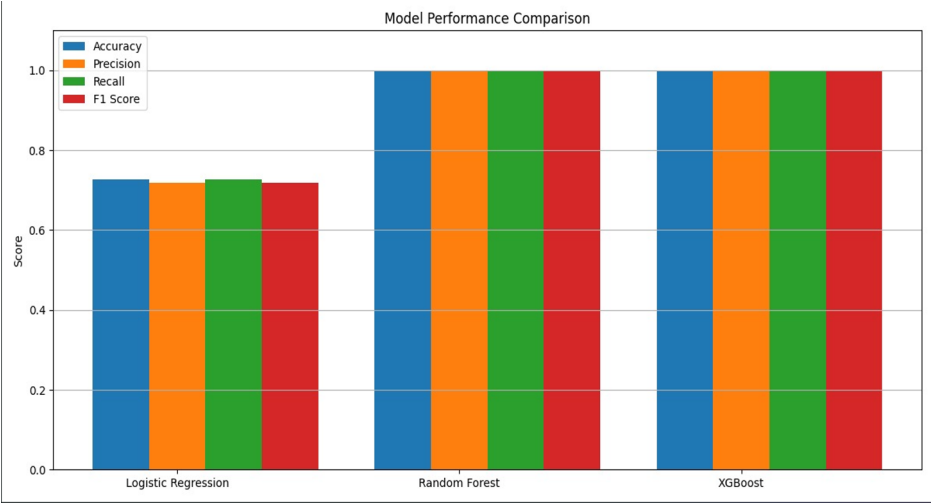
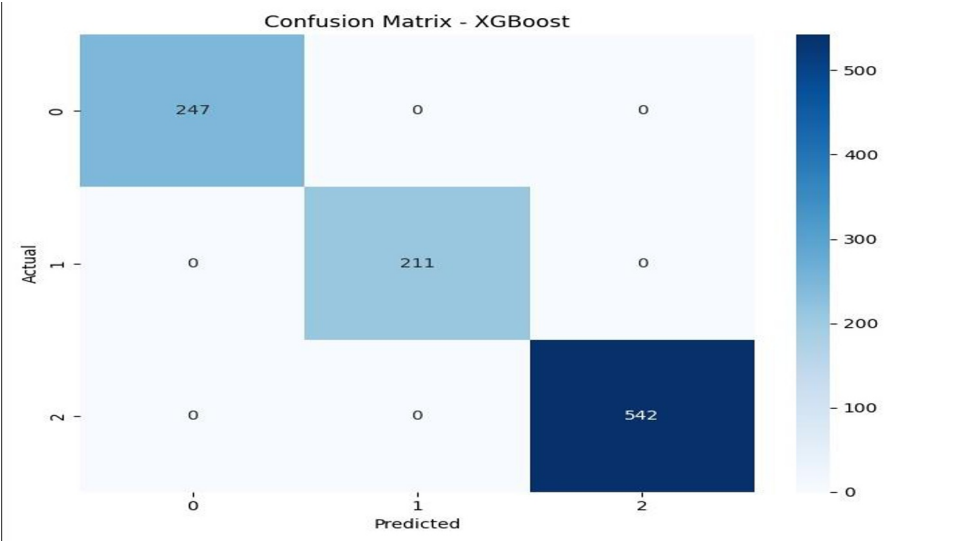


Fig. 2. Classifier Performance Comparison

Confusion matrix was also plotted to analyse class-wise prediction outcome. Fig.3 demonstrates the detailed class-wise evaluation.



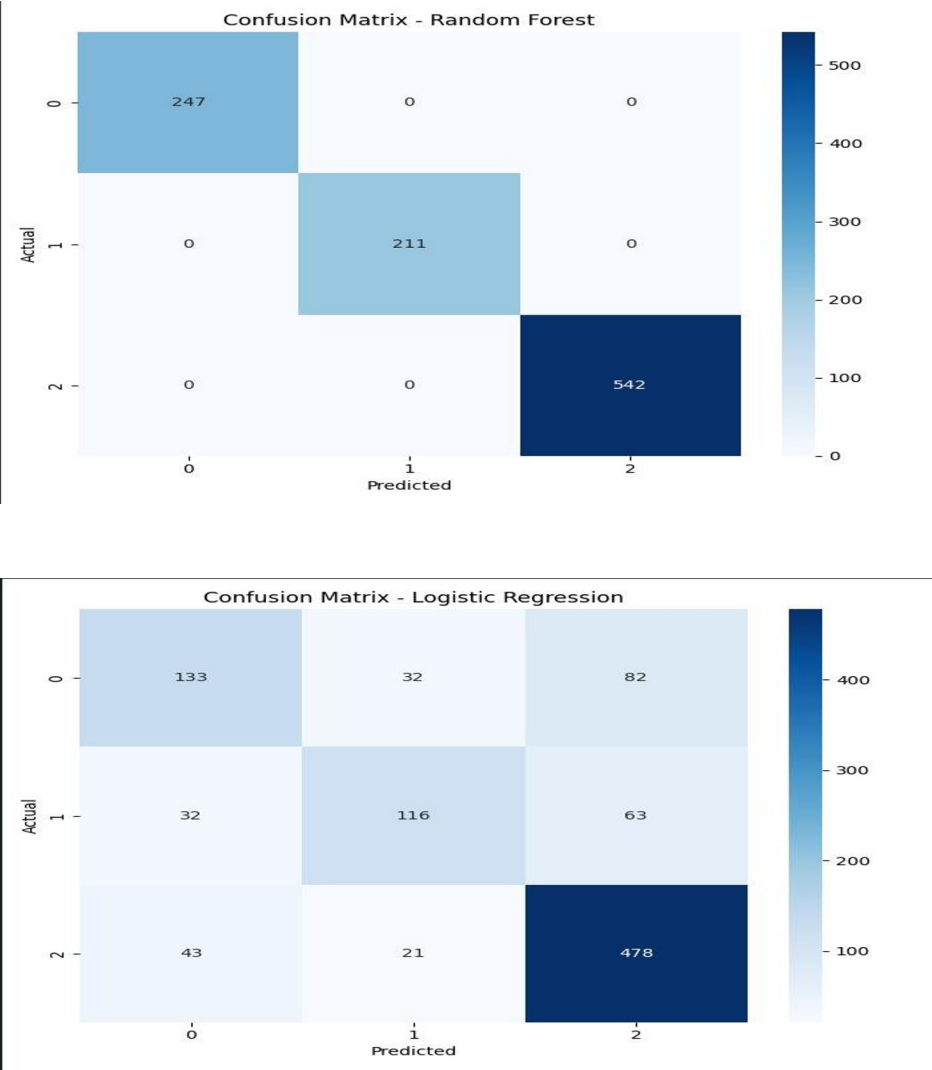


Fig. 3. Confusion matrix for ML classifiers under consideration

Table 2 presents the comparative performance of all three ML models, including class-wise Precision, Recall, and F1-Score.

Table 2. ML model performance

Model	Accuracy	Precision	Recall	F1-Score
-------	----------	-----------	--------	----------

Logistic Regression	0.81	0.81	0.81	0.81
Random Forest	1.00	1.00	1.00	1.00
XGBoost	1.00	1.00	1.00	1.00

The accuracy results for Random Forest and XGBoost from Table 2 needs appropriate understanding model’s behaviour at training and testing. So in line with this diagnosis of accuracy it has been observed that the dataset curated was mapping of dosha and symptom which is very deterministic. The relationship between feature and label is highly structured and internally consistent. Ensemble models such as Random Forest and XGBoost learns such structured patterns with high precision. Additionally, to avoid over fitting model was trained with 5-fold cross validation. However the cross-validation also resulted in accuracy value as 1.00 across all folds with zero standard deviation. This ensures that model’s learning was consistent and not getting over-fitted. These results conclude that synthetic data has deterministic nature and hence it may not be the indicative of performance for real clinical populations. As this is proof of concept level work real-world clinical data validation and integration is a necessary prerequisite before any deployment consideration. A confusion matrix for the Random Forest model of Fig. 3 confirms perfect separation across all three Dosha classes with no misclassifications. Random Forest was selected as the primary model for the live recommendation pipeline due to its robustness, feature importance interpretability, and stable generalization behaviour.

5.2 Retrieval Performance

The FAISS IndexFlatL2 retrieval system provided sub-millisecond latency per query on the Ayurvedic knowledge base, allowing fast semantic searches. It has been found that semantic relatedness of therapies and herb records were consistently prioritized irrespective of precise keyword matching, retrieval handled extremely well with respect to variations in symptom descriptions as well as Ayurvedic terminology, contextual match was enormously superior than a keyword based lookup which reflected on the downstream relevance of recommendation.

The retrieval capability of the proposed RAG framework is validated by analysing representative retrieval examples. Total 50 diverse migraine-related queries which include 10 queries for each Dosha along-with 10 queries based on symptoms related migraine0 and 10 queries based on herbal remedies were used for evaluating the performance of retrieved results from the FAISS vector database. Table 3 exhibits the results of sample records retrieved. The retrieved chunks clearly indicate not only semantic alignment with user intent of Dosha characteristics but also returned therapy recommendations to address this Dosha imbalance. Additionally, symptom-oriented queries like photophobia, nausea, vomiting successfully recommended migraine-specific therapy. So it is clearly evident that the relevant Ayurvedic treatment records are successfully retrieved using the embedding-based retrieval mechanism and thereby providing a strong foundation for a reliable knowledge base for subsequent recommendation generation.

Table 3. Sample Retrieved Records from the FAISS-Based RAG System

Query	Retrieved Chunk ID	Source Document	Retrieved Record
Vata dosha migraine throbbing headache	TM-0066	therapy_materials.csv	Therapy: Virechana; Material: Milk; Targets Dosha: Vata; Headache Type: Migraine; Benefits: Balances Vata dosha and reduces headache symptoms.
Vata dosha migraine throbbing headache	TM-0007	therapy_materials.csv	Therapy: Nasya Karma; Material: Ghee; Targets Dosha: Vata; Headache Type: Migraine; Benefits: Reduces headache symptoms associated with Vata imbalance.
Pitta dosha burning headache stress heat migraine	TM-0229	therapy_materials.csv	Therapy: Upanaha; Material: Cooling Ghee; Targets Dosha: Pitta; Headache Type: Migraine; Benefits: Balances Pitta dosha and alleviates burning headache symptoms.
Pitta dosha burning headache stress heat migraine	TM-0100	therapy_materials.csv	Pitta-oriented migraine therapy recommendation with cooling therapeutic material and dosha- balancing properties.
Kapha dosha dull headache congestion sinus heaviness	TM-0551	therapy_materials.csv	Therapy: Nasya Karma; Material optimized for Kapha-related migraine symptoms including congestion and heaviness.
Kapha dosha dull headache congestion sinus heaviness	TM-0541	therapy_materials.csv	Therapy: Nasya Karma; Material: Haridra Smoke (Winter, Moderate Dose); Targets Dosha: Kapha; Benefits: Reduces Kapha-related headache symptoms.
Kapha dosha dull headache congestion sinus heaviness	TM-0542	therapy_materials.csv	Kapha-specific migraine treatment record with seasonal optimization and dosage recommendations.

Query	Retrieved Chunk ID	Source Document	Retrieved Record
Migraine with nausea, vomiting, sensitivity to light	TM-0034	therapy_materials.csv	Therapy: Shirodhara; Material: Bala Taila; Targets Dosha: Vata; Benefits: Balances Vata dosha and reduces migraine symptoms.
Migraine with nausea, vomiting, sensitivity to light	TM-0607	therapy_materials.csv	Therapy: Shirodhara; Material: Bala Taila (Summer, Moderate Dose); Optimized for Summer season; Migraine-specific treatment.
Ayurvedic herbs for headache relief natural remedy	TM-0505	therapy_materials.csv	Therapy: Nasya Karma; Material: Ginger Paste (Winter, Moderate Dose); Targets Dosha: Kapha; Benefits: Reduces headache symptoms.
Ayurvedic herbs for headache relief natural remedy	TM-0502	therapy_materials.csv	Therapy: Nasya Karma; Material: Ginger Paste (Monsoon, Moderate Dose); Seasonal herbal remedy for migraine management.

5.3 LLM-Based Recommendation Quality

The LLaMA3 model generated personalized Ayurvedic recommendations in a clear, natural, and easy-to-understand manner. The system gives its suggestions based on patient’s dosha type. For patients with Vata imbalance, it recommended warming and calming therapies. For Pitta imbalance, it focused on cooling diets and avoids heat-producing food. For Kapha imbalance, it suggested detoxification methods. All these recommendations were compared with classical Ayurvedic texts and were found to be accurate. Dosha prediction and therapy report is shown in Table 4.

Table 4.Retrieval Performance Metrics of the Proposed RAG System

Query Category	Top-1 Score	Avg Score	Top-3 Avg Score	Top-5
Vata Dosha Migraine	0.5832	0.5798	0.5770	
Pitta Dosha Migraine	0.6974	0.6962	0.6955	
Kapha Dosha Migraine	0.6389	0.6380	0.6372	

Query Category	Top-1 Score	Avg Score	Top-3 Avg Score	Top-5
Migraine Symptoms (Nausea, Photophobia)	0.3109	0.3077	0.3051	
Ayurvedic Herbal Remedies	0.5751	0.5728	0.5716	

The retrieval effectiveness has been assessed with Top-5 retrieval results for 50 representative migraine-related queries. As per the results shown in Table 4, it has been observed that the proposed system has returned relevant therapy records for dosha-specific queries with a cosine similarity scores ranging from 0.57–0.70. Also, by analysing every Dosha it is evident that Pitta-related migraine queries have highest retrieval performance with Top-1 similarity as 0.6974, which clearly indicates that it has a strong semantic relation between query embedding and stored therapeutic records. Also, retrieval stability and consistency is validated with Top-1 and Top-5 average similarity score. The system is very consistent as there is a very small difference between these scores. So it can be concluded that the proposed retrieval system is suited for building contextually relevant knowledgebase for Ayurveda.

Table 5. Average Retrieval Performance Across All Queries

Metric	Value
Average Precision@1	1.0000
Average Recall@1	0.0020
Average Precision@3	1.0000
Average Recall@3	0.0059
Average Precision@5	1.0000
Average Recall@5	0.0098

Table 6. Cosine Similarity Analysis

Query	Avg Top-1 Similarity	Avg Top-3 Similarity	Avg Top-5 Similarity
Vata dosha migraine throbbing headache	0.5832	0.5798	0.5770
Pitta dosha burning headache stress heat migraine	0.6974	0.6962	0.6955
Kapha dosha dull headache congestion sinus heaviness	0.6389	0.6380	0.6372

Query	Avg Top-1 Similarity	Avg Top-3 Similarity	Avg Top-5 Similarity
Migraine with nausea, vomiting, sensitivity to light	0.3109	0.3077	0.3051
Ayurvedic herbs for headache relief natural remedy	0.5751	0.5728	0.5716

Table 7. Overall Similarity Statistics

Statistic	Value
Mean Cosine Similarity	0.5573
Maximum Cosine Similarity	0.6974
Minimum Cosine Similarity	0.3008

Precision@K and Recall@K metrics are used to find the effectiveness of the proposed recommender. As observed in Table 5 Precision@1, Precision@3, and Precision@5 resulted in a value of 1.000 for all the queries under consideration, indicating the relevance of query with the retrieved result. Similarly, as ‘k’ value increases the Recall values are also increased consistently indicating improved relevant records coverage. Cosine similarity analysis is carried out for validating semantic retrieval quality. Similarity scores are ranging from 0.3051 to 0.6955 for different query categories and exhibited in Table 6. Overall similarity statistics is exhibited in Table 7. with an overall mean cosine similarity of 0.5573 and a maximum similarity of 0.6974All these results confirms that the proposed system has effectively used retrieval module for retrieving semantic aware Ayurvedic therapy recommendations

6 Future Scope

The existing work can be expanded in future in many dimensions. One of the major concerns is integrating real time clinical data under clinical expert’s guidance along with appropriate validation. With the current framework it is actually well understood about various features required to build this framework. Currently only Tridoshas are under considerations but it can be extended with more features like voice analysis, facial features wearable sensor data etc. With these features a perfect multimodal system architecture is quite possible for better recommendations. Also, specialized healthcare Transformer based models like Med-PaLM and BioBERT has to be tested for validating quality of solutions. As data increases definitely better indexing strategies need to be incorporated so in that case advanced indexing like Hierarchical Navigable Small World graphs may scale up retrieval without slowing down. Constructive feedbacks always makes system better so in this system too techniques like Reinforcement learning can be integrated for more realistic recommendations. Deployment of this system as a app will have a great impact on digital footprints.

7 Conclusion

The proposed work is an attempt to develop AI-based framework acting as a bridge between ancient knowledge equipped with contemporary techniques to address a major neurological disorder. The proposed system makes use of ML classifiers for predicting Doshas, interacting with users to understand the symptoms through RAG based framework, encoding the ancient knowledge into technological understanding with sentence transformer embedding, retrieving data with FAISS and lastly generating the responses with LLMs specially LLaMA3 for the current system

The proposed work provides a potential leap for modern and personalized healthcare systems.

Acknowledgments. The authors would like to thank the management and administration of our institute for providing the necessary facilities and support to carry out this research work.

Disclosure of Interests. The authors declare that they have no competing interests.

References

1. Soman, A., Venkatram, S., Chikkanna, U., Ramakrishna, K.K., Bhargav, H., Mailankody, P., Varambally, S.: Ayurveda for management of migraine: A narrative review of clinical evidence. *J. Fam. Med. Prim. Care* 11(8), 4228–4235 (2022). https://doi.org/10.4103/jfmpe.jfmpe_2109_21
2. Singh, S., Ashu, Thakur, P.: Evaluating the therapeutic potential of Ayurvedic interventions in migraine: A review of experimental and clinical evidence. *Int. J. AYUSH* 14(10) (2024). <https://doi.org/10.22159/prl.ijayush.v14i10.1563>
3. Shaikh Chand, S.A., Naaz, S.A., Junaid, P.M., Khan, T.K.: Nasya therapy in the management of Ardhavabhedaka (migraine): A review of Ayurvedic perspectives and modern insights. *Afr. J. Biomed. Res.* 27(4S) (2024). <https://doi.org/10.53555/AJBR.v27i4S.5326>
4. Amugongo, L.M., Mascheroni, P., Brooks, S.: Retrieval Augmented Generation for Large Language Models in Healthcare. *PLOS Digital Health* (2025)
5. Liu, S., McCoy, A.B., Wright, A.: Improving large language model applications in biomedicine with retrieval-augmented generation. *J. Am. Med. Inform. Assoc.* (2025)
6. World Health Organization: Neurological Disorders: Public Health Challenges. World Health Organization, Geneva (2006). ISBN 9789241563369
7. Rani, B., et al.: Clinical study on Pathyadi Kwatha and Nasya therapy in migraine. *AYUSHDHARA* (2024)

8. Navyashree, N.M.S., et al.: Comparative clinical study of Nasya Karma and Shirodhara in migraine. *J. Ayurveda Integr. Med. Sci.* (2020)
9. Yarnell, E.: Herbal medicine and migraine. *Altern. Complement. Ther.* 23(5) (2017)
10. Kanakhara, B., Chaudhari, V.: Clinical efficacy of Agnikarma and Pathyadi decoction in migraine. *AYU* (2018)
11. Sharma, S., et al.: Role of Panchakarma in migraine management. *J. Ayurveda* (2018)
12. Pedregosa, F., Varoquaux, G., Gramfort, A., et al.: Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* 12, 2825–2830 (2011)
13. Chen, T., Guestrin, C.: XGBoost: A scalable tree boosting system. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794 (2016)
14. Devlin, J., Chang, M.-W., Lee, K., Toutanova, K.: BERT: Pre-training of deep bidirectional transformers for language understanding. In: *Proceedings of NAACL-HLT*, pp. 4171–4186 (2019)
15. Ng, K.K.Y., Matsuba, I., Zhang, P.: RAG in healthcare: A novel framework for improving communication and decision-making. *Inf. Fusion* (2026). <https://doi.org/10.1016/j.inffus.2026.104483>
16. Miao, Y., Zhao, Y., Luo, Y., Wang, H., Wu, Y.: Improving large language model applications in the medical and nursing domains with retrieval-augmented generation: Scoping review. *J. Med. Internet Res.* 27, e80557 (2025). <https://doi.org/10.2196/80557>
17. Kuo, S.M., et al.: Automated clinical trial data analysis and report generation by integrating retrieval-augmented generation (RAG) and large language model (LLM) technologies. *AI* 6(8), 188 (2025). <https://doi.org/10.3390/ai6080188>
18. Balakrishnan, P., et al.: Safety evaluation of Nasya therapy. (2023)
19. Gupta, R., et al.: Comparative analysis of Ayurveda and allopathy in migraine treatment. *J. Integr. Med.* (2022)
20. Singh, S.K., et al.: Evaluation of Nasya therapy in migraine management: A randomized clinical trial. *AYU Journal* (2023)
21. Patil, G., Patil, S., Hosur, P.: Ayurveda management of migraine – a case report. *J. Ayurveda Integr. Med.* 15(5), 100983 <https://doi.org/10.1016/j.jaim.2024.100983>
22. Abo El-Enen, M., Saad, S., Nazmy, T.: A survey on retrieval-augmentation generation (RAG) models for healthcare applications. *Neural Comput. Appl.* 37(33), 28191–28267 (2025). <https://doi.org/10.1007/s00521-025-11666-9>
23. Vaghela D., Shweta M., Dhiman K., Manjusha R BA randomized clinical trial on Ardhavabhedaka–migraine and its Ayurvedic management. *Int. J. Ayurvedic Med.* 7(1) (2016). <https://doi.org/10.47552/ijam.v7i1.721>
24. Charaka Samhita Online. https://www.carakasamhitaonline.com/index.php?title=Main_Page. Accessed 10 Jun 2026
25. Sushruta Samhita (NIIMH Digital Edition). <https://niimh.nic.in/ebooks/esushruta/?mod=read>. Accessed 10 Jun 2026

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

