



Multi-Label Machine Learning Framework for Homoeopathic Remedy Recommendation using Clinical Disease Descriptions

Rincy T.A., PhD
Associate Professor
Prajyoti Niketan College, Pudukad
Thrissur, Kerala

ABSTRACT

The application of artificial intelligence in healthcare has created new opportunities for developing data-driven clinical decision-support systems. This study explores the use of machine learning for recommending multiple Homoeopathic remedies based on clinical disease descriptions. A multi-label classification framework was developed using text preprocessing and TF-IDF-based feature representation. Different machine learning algorithms were comparatively evaluated to identify their effectiveness in predicting associated remedies. The findings demonstrate the potential of machine learning techniques for automated remedy recommendation, while also highlighting challenges related to data sparsity, class imbalance, and the diversity of remedy associations. The proposed approach provides a foundation for developing intelligent tools that can support practitioners in analysing clinical descriptions and identifying potential remedy options.

General Terms

Homoeopathy, Artificial Intelligence, Complementary Medicine.

Keywords

Machine Learning, Homoeopathic Remedy Recommendation, Multi-label Classification, TF-IDF, Natural Language Processing.

1. INTRODUCTION

Artificial Intelligence (AI) has emerged as a transformative technology in modern healthcare, enabling the development of intelligent systems for disease diagnosis, treatment planning, prognosis prediction and clinical decision support. Machine learning (ML) and Natural Language Processing (NLP) techniques have shown great success in extracting meaningful information from medical records, clinical narratives and healthcare databases, helping healthcare professionals to make more accurate and evidence-informed decisions. The accuracy, consistency and efficiency of healthcare decisions have been improved by Machine Learning technologies, making them integral components of many healthcare applications.

Homoeopathy is an alternate system of medicine in which the principle of individualization, playing key role in the prescription of medicine. As the aforementioned system considers the disease of the patient, its symptoms, pathology, as well as the individual peculiarities of the patient, the knowledge in *Materia Medica* and *Repertories* as well as clinical know-how of the doctor plays a vital role in the derivation and

prescription of exact medicine. As there are many medicines which may be considered in every clinical condition, pinpointing to a single medicine is a tedious task. Hence, a prevailing need can be identified among homoeopaths for an analytical tool that assists in the selection of appropriate medicines while ensuring that clinical assessment remains uncompromised.

Recent advances in machine learning have enabled the development of intelligent clinical decision-support systems that assist healthcare practitioners by analyzing medical data and recommending potential treatments. Although artificial intelligence has been extensively applied in conventional medicine for disease diagnosis and treatment recommendation, research on automated homoeopathic remedy recommendation remains limited. Most existing studies address conceptual frameworks, expert systems, or literature reviews, while few evaluate modern machine learning algorithms using curated homoeopathic prescription datasets. Also the challenges regarding much more typical remedy distribution of Homoeopathic medicines, have also received limited attention.

The finding of a Homoeopathic *simillimum* involves peculiar and multiple challenges for computation. First, a single disease can correspond to multiple remedies, which renders the problem a multi-label classification task rather than a traditional single-label classification problem. Second, remedy frequencies are highly imbalanced, with a small subset of remedies prescribed frequently and many others occurring rarely. Third, disease descriptions are generally brief textual phrases that necessitate the use of natural language processing techniques to convert them into meaningful numerical representations for machine learning algorithms.

To address these challenges, a machine learning-based framework is proposed for multi-label homoeopathic remedy recommendation based on disease descriptions. A curated dataset comprising approximately 1,300 disease records and 364 standardized homoeopathic remedies was prepared through systematic preprocessing. Disease descriptions were cleaned using established natural language processing (NLP) techniques, and remedy names were standardized by removing potency, dosage, and dispensing information to ensure consistent labeling. The processed disease nomenclatures were then strategically transformed into numerical feature vectors using the Term Frequency–Inverse Document Frequency (TF-IDF) technique.

Multiple supervised machine learning algorithms, including Logistic Regression, Linear Support Vector Machine (SVM),



Random Forest, and Extreme Gradient Boosting (XGBoost), were implemented and compared for multi-label remedy prediction. Model performance was evaluated using Micro-F1 Score, Macro-F1 Score, Precision, Recall, Jaccard Similarity, and Hamming Loss to provide a comprehensive assessment of predictive performance under highly imbalanced multi-label conditions.

The outcomes of this research demonstrate the feasibility of applying machine learning techniques to homeopathic remedy recommendation and contribute toward the development of intelligent clinical decision-support systems for complementary medicine. The evolved framework also furnishes ample opportunity for future research involving larger datasets, contextual language models, explainable artificial intelligence, as well as extensive clinical validation. Furthermore, integrating patient-specific clinical features and longitudinal health records could improve the personalization and reliability of remedy recommendations. Such advancements may enhance the practical applicability of the proposed system and support evidence-based decision-making in complementary medicine.

2. LITERATURE REVIEW

Artificial Intelligence (AI) has emerged as a transformative technology in healthcare by improving disease diagnosis, clinical decision support, personalized treatment, and predictive analytics. AI-based systems assist healthcare professionals in analyzing large volumes of clinical data and generating evidence-based recommendations, thereby improving the quality and efficiency of patient care [1]. Recent advancements in machine learning have further enabled the development of intelligent decision-support systems capable of processing complex medical information and assisting clinicians in treatment planning.

The integration of digital technologies into Homoeopathic practice has received increasing attention in recent years. Digital case-taking systems, electronic health records, and AI-assisted clinical applications have been introduced to improve data management, treatment consistency, and practitioner efficiency [2,3,4]. These technologies facilitate systematic documentation of patient information while reducing manual errors and improving the accessibility of clinical records [3,5].

Clinical Decision Support Systems (CDSS) assist healthcare professionals by providing evidence-based recommendations during diagnosis and treatment planning. Because they can learn complex relationships from clinical data, traditional machine learning algorithms like Logistic Regression, Support Vector Machines (SVM), Random Forests, and Extreme Gradient Boosting (XGBoost), Clinical BERT have shown excellent performance in a variety of healthcare prediction tasks [6,7,9].

The success of AI-based clinical decision-support systems largely depends on the availability of high-quality clinical data. Standardized data collection protocols and structured electronic case records improve model training, enhance reproducibility, and facilitate the development of reliable predictive models [10–13]. Researchers have emphasized that data quality, consistency, and standardization are essential prerequisites for the successful implementation of machine learning in complementary medicine [14,15]. These studies suggest that carefully curated datasets contribute significantly to improving model performance and clinical reliability.

Relatively few studies have evaluated supervised machine

learning algorithms using curated Homoeopathic prescription datasets. Some researchers have investigated the application of machine learning specifically in Homoeopathy. Pattern recognition techniques have been employed to develop intelligent systems capable of identifying constitutional characteristics and recommending appropriate remedies based on patient symptoms [8,14,15]. Deep learning approaches have also been explored for constitutional analysis, demonstrating the potential of neural networks to capture complex relationships between clinical features and remedy selection [16]. Furthermore, predictive modeling has been successfully applied to estimate treatment outcomes, highlighting the growing role of AI in supporting Homoeopathic clinical practice [17].

Limited attention has been given to modelling Homoeopathic remedy recommendation as a multi-label text classification problem, where multiple remedies may correspond to a single disease description. This represents an important research gap in the existing literature. Although previous research has demonstrated the feasibility of applying machine learning in Homoeopathy, relatively few studies have investigated multi-label Homoeopathic remedy recommendation using disease descriptions as textual input. Most existing approaches concentrate on constitutional analysis, treatment outcome prediction, or practitioner support rather than predicting multiple remedies simultaneously [14,15,16,17].

Natural Language Processing (NLP) plays a significant role in evaluating unstructured clinical information such as electronic health records, discharge summaries, physician notes, and disease descriptions. NLP techniques including tokenization, lemmatization, stop-word removal, and feature extraction enable machine learning algorithms to transform textual information into meaningful numerical representations.

Several challenges remain in applying AI to Homoeopathy. One of the major limitations reported in previous studies is the scarcity of standardized, large-scale datasets suitable for machine learning applications [18]. In addition, issues such as class imbalance, limited sample sizes, heterogeneous clinical records, and variability in practitioner prescribing patterns continue to affect model performance and generalizability [19,20].

The present study develops and evaluates a multi-label machine learning framework for recommending multiple Homoeopathic remedies from disease descriptions. The framework employs TF-IDF for textual feature extraction and evaluates multiple supervised learning algorithms, including Logistic Regression, Linear Support Vector Machine, Random Forest, and XGBoost. A comparative evaluation of these classifiers is conducted to examine their effectiveness in automated Homoeopathic remedy recommendation. The study further contributes to the development of AI-assisted approaches for supporting remedy selection based on textual clinical descriptions.

3. PROPOSED FRAMEWORK

3.1 Overview

This section introduces the overall architecture of the proposed Homoeopathic remedy recommendation system. The framework is designed to recommend multiple Homoeopathic remedies from disease descriptions using Natural Language Processing (NLP) and supervised machine learning techniques. Since a single disease may correspond to multiple remedies, the recommendation problem is formulated as a multi-label text

classification task.

The proposed framework consists of six sequential stages: dataset collection, data preprocessing, multi-label encoding, feature extraction, machine learning-based remedy prediction, and performance evaluation. Logistic Regression, Linear Support Vector Machine (SVM), Random Forest, and Extreme Gradient Boosting (XGBoost) are the Machine Learning models used to train the data. The models are evaluated using standard multi-label classification metrics, including Micro-F1 Score, Macro-F1 Score, Precision, Recall, Jaccard Similarity, and Hamming Loss.

The overall workflow of the proposed framework is illustrated in Figure 1.

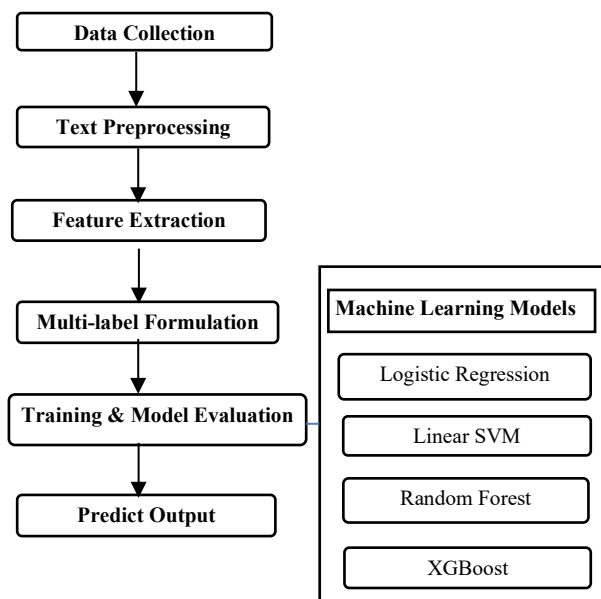


Fig 1: Homoeopathic Remedy Recommendation System

3.2 Methodology

A curated dataset containing approximately 1290 disease records and 364 standardized Homoeopathic remedies was prepared for this study. Each record consists of a disease description and its associated Homoeopathic remedies. Since multiple remedies may correspond to a single disease, the dataset represents a multi-label classification problem.

The remedy names were standardized by removing potency values (e.g., 30C, 200C, 1M), dosage quantities, pharmaceutical forms (powder, tablets, pills), and other dispensing information. This preprocessing reduced label duplication and ensured consistency across the dataset.

Disease descriptions were preprocessed using Natural Language Processing techniques to improve data quality. The preprocessing pipeline consisted of Conversion to lowercase, Removal of punctuation, numbers, and special characters, Tokenization, Lemmatization using the WordNet Lemmatizer, Removal of unnecessary textual variations.

The preprocessing stage reduced textual noise while preserving clinically meaningful information. The standardized remedies associated with each disease were converted into binary vectors using the MultiLabelBinarizer technique. Each remedy represents an independent output label.

The processed disease descriptions were converted into numerical feature vectors using the Term Frequency–Inverse

Document Frequency (TF-IDF) technique. TF-IDF assigns weights to words according to their importance within the dataset while reducing the influence of highly frequent terms. The TF-IDF vectorizer was configured with a maximum vocabulary size, generating a sparse numerical representation suitable for text classification.

Four supervised machine learning algorithms were implemented:

- Logistic Regression

Logistic Regression serves as a baseline linear classifier for multi-label text classification.

- Linear Support Vector Machine

Linear SVM constructs optimal separating hyperplanes in high-dimensional feature spaces and is widely recognized for text classification tasks.

- Random Forest

Random Forest is an ensemble learning algorithm that combines multiple decision trees to improve prediction accuracy and reduce overfitting.

- Extreme Gradient Boosting (XGBoost)

XGBoost is a gradient boosting algorithm capable of modelling complex non-linear relationships while incorporating regularization mechanisms to reduce overfitting. Each algorithm was implemented using the One-vs-Rest (OvR) strategy, where an independent binary classifier was trained for each remedy class.

The dataset was divided into training and testing subsets using an 80:20 train-test split. The performance of each model was assessed using standard multi-label classification metrics: Micro-F1 Score, Macro-F1 Score, Precision, Recall, Jaccard Similarity, and Hamming Loss. These metrics provide a comprehensive evaluation of prediction performance under class imbalance and multi-label conditions.

4. RESULTS AND DISCUSSION

This study evaluated the performance of four machine learning models for multi-label Homoeopathic remedy recommendation based on disease descriptions. The experimental results revealed noticeable differences in the predictive capabilities of the evaluated models, indicating that the choice of classifier significantly influences recommendation performance.

Among the four models, the TF-IDF + Random Forest classifier achieved the best overall performance, obtaining the highest Micro-F1 score, Recall, and Jaccard Similarity. These results suggest that Random Forest was more effective in identifying the correct combinations of remedies for each disease. The ensemble learning mechanism of Random Forest enables it to model complex and non-linear relationships among TF-IDF features, which is advantageous when multiple remedies are associated with a single disease. Its higher recall indicates that the model successfully retrieved more relevant remedies than the other classifiers, making it more suitable for recommendation tasks where missing relevant remedies is undesirable.

Table 1. Micro F1, Macro F1 and Precision Scores

Model	Micro F1	Macro F1	Precision
TF-IDF + Logistic Regression	0.1387	0.0057	0.5573
TF-IDF + Linear SVM	0.2051	0.0325	0.4690
TF-IDF + Random	0.2146	0.0316	0.4381

Forest			
TF-IDF + XGBoost	0.1791	0.0147	0.4700

The TF-IDF + Linear SVM model demonstrated competitive performance, achieving the second-highest Micro-F1 score and the highest Macro-F1 score. Linear SVM is well suited to high-dimensional sparse text representations such as TF-IDF, allowing it to effectively discriminate between remedy classes. The comparatively higher Macro-F1 score suggests that SVM provided more balanced performance across remedy classes than the other models, although the overall score remained low because of severe class imbalance within the dataset.

Table 2. Recall, Jaccard Similarity and Hamming Loss

Model	Recall	Jaccard Similarity	Hamming Loss
TF-IDF + Logistic Regression	0.0792	0.0745	0.00966
TF-IDF + Linear SVM	0.1312	0.1143	0.00999
TF-IDF + Random Forest	0.1421	0.1202	0.01021
TF-IDF + XGBoost	0.1106	0.0984	0.00996

The TF-IDF + XGBoost model achieved moderate performance, with a Micro-F1 score of 0.1791. Although it produced high precision, its comparatively lower recall reduced its overall effectiveness. While XGBoost has demonstrated excellent performance in many classification tasks, its advantages were not fully realized in this study. It may be due to the relatively small and highly imbalanced dataset, combined with sparse TF-IDF features. These characteristics limit the ability of gradient boosting algorithms to learn reliable decision boundaries for infrequently occurring remedy classes.

The TF-IDF + Logistic Regression model achieved the highest precision and the lowest Hamming Loss, indicating that its predictions were relatively accurate when a remedy was recommended. However, its low recall and the lowest Micro-F1 score reveal that the model produced conservative predictions, recommending fewer remedies and consequently missing many relevant labels. Although Logistic Regression serves as a strong baseline for text classification, it was less effective than Random Forest and Linear SVM for the proposed multi-label recommendation task.

An important factor influencing the performance of all models is the class imbalance within the dataset. The dataset consists of 1,290 disease records mapped to 364 unique remedies, of which 101 remedies occur only once and 153 remedies occur fewer than three times. Such imbalance makes it difficult for the models to learn reliable representations for rare remedies. This explains the consistently low Macro-F1 scores across all classifiers, whereas the comparatively higher Micro-F1 scores indicate better performance on frequently occurring remedies. Therefore, the observed performance reflects not only the strengths of the classifiers but also the inherent complexity of the dataset. Fig 1 and Fig 2 shows the Comparative performance of TF-IDF-based classifiers for multi-label classification.

Model Performance: F1 Score, Precision and Recall

Comparison of TF-IDF-based machine learning models across three classification metrics.

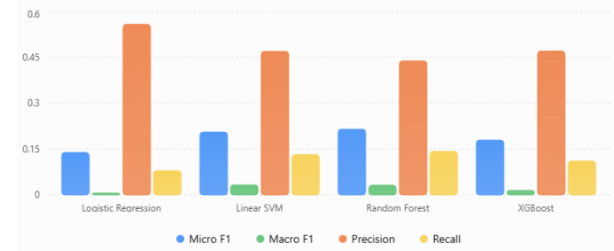


Fig 1: Model Performance

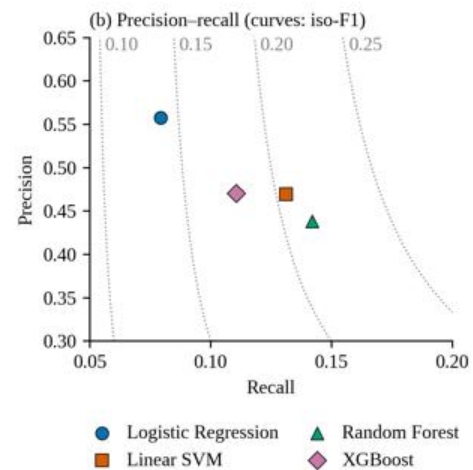


Fig 2: Precision-recall scatter plot with iso-F1 contours of equal micro-averaged F1.

The results also demonstrate that traditional machine learning approaches remain effective for sparse medical text classification. However, the performance achieved suggests that further improvements are possible through larger and more balanced datasets, hyperparameter optimization, advanced text representations such as biomedical language models, and techniques specifically designed for imbalanced multi-label learning. These enhancements represent promising directions for future research and could further improve the accuracy and clinical applicability of automated Homoeopathic remedy recommendation systems.

5. CONCLUSION

This study introduced a multi-label machine learning framework for recommending Homoeopathic remedies based on clinical disease descriptions. The textual data were represented using term frequency-inverse document frequency (TF-IDF), and four machine learning models—Logistic Regression, Linear Support Vector Machine (SVM), Random Forest, and XGBoost—were evaluated. Random Forest demonstrated the highest overall performance, achieving superior Micro-F1, Recall, and Jaccard Similarity scores. These findings support the feasibility of applying machine learning to multi-label Homoeopathic remedy recommendation. Nevertheless, the relatively low Macro-F1 scores highlight challenges associated with dataset imbalance and limited representation of rare remedies. Future research should prioritize expanding and balancing the dataset, optimizing model parameters, and investigating advanced text representations and deep learning methods to enhance



recommendation performance and generalizability.

6. REFERENCES

- [1] Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H. and Wang, Y. 2017 Artificial intelligence in healthcare: Past, present and future. *Stroke and Vascular Neurology*. 2(4): 230-243.
- [2] Banerjee, K. et al. 2020 Digital health interventions in homeopathy: A systematic review. *Homeopathy*. 109(04):189-198.
- [3] Manchanda, R.K. et al. 2022 Digital transformation in homeopathic practice: A systematic review of technological interventions. *Homeopathy*. 111(02):108-117.
- [4] Wu, H. and Thompson, M. 2023 Integration of digital health technologies in complementary medicine practice. *Journal of Medical Systems*. 47(1):15.
- [5] Thompson, R. et al. 2023 Clinical decision support systems in homeopathy: Current status and future perspectives. *Homeopathy*. 112(01):15-24.
- [6] Bera, M.N. and Gupta, A. 2025 Homoeopathy Meets Machine Learning: A Systematic Review of Pattern Recognition in Patient Constitutional Types.
- [7] Alsentzer, E., Murphy, J.R., Boag, W., Weng, W.H., Jin, D., Naumann, T. and McDermott, M. 2019 Publicly Available Clinical BERT Embeddings. *Proceedings of the 2nd Clinical Natural Language Processing Workshop*.
- [8] Park, S.H. et al. 2022 Digital case-taking systems in homeopathy: Impact on data quality and clinical outcomes. *BMC Complementary Medicine and Therapies*. 22(1):156
- [9] Chen, X. et al. 2023 Standardization protocols for machine learning applications in complementary medicine. *Journal of Medical Systems*. 47(2):28.
- [10] Anderson, K. et al. 2022 Framework development for digital data collection in homeopathic practice. *BMC Complementary Medicine and Therapies*. 22(2):178.
- [11] Sharma, R. and Wilson, J. 2023 Quality control in digital homeopathy: A systematic approach. *Digital Health*. 9:20552076231158392.
- [12] Rossi, E. et al. 2021 Machine learning applications in constitutional homeopathy: Development and validation of a pattern recognition system. *Digital Health*. 7:20552076211006948.
- [13] Martinez, A. et al. 2023 Supervised learning in homeopathic prescribing: A validation study. *Artificial Intelligence in Medicine*. 135:102509.
- [14] Kim, S. et al. 2022 Deep learning applications in constitutional homeopathy. *Journal of Alternative and Complementary Medicine*. 28(8):789-798.
- [15] Lee, J.Y. et al. 2023 Predictive modeling in homeopathic treatment outcomes: A machine learning approach. *Complementary Therapies in Medicine*. 70:102928.
- [16] Breiman, L. 2001 Random Forests. *Machine Learning*. 45(1):5-32.
- [17] Chen, T. and Guestrin, C. 2016 XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.
- [18] Williams, M. et al. 2023 Machine learning in homeopathy: A systematic review and meta-analysis. *Digital Health*. 9:20552076231159876.
- [19] Davidson, R. and Park, S. 2023 Implementation challenges in digital homeopathy. *Journal of Medical Systems*. 47(3):45.
- [20] Liu, J. et al. 2022 Technical barriers in implementing AI systems in complementary medicine. *Artificial Intelligence in Medicine*. 134:102447.