Research Journal of Cell Sciences

PRIME OPEN ACCESS

Volume 1, Issue 1

Research Article

Date of Submission: 01 April, 2025
Date of Acceptance: 30 April, 2025
Date of Publication: 06 May, 2025

Multimodal Decision Support System for Improved Diagnosis and Healthcare Decision Making

Aniket Patil*, Viraj Patil, Sangram Sankpal, Tanuja S. Patankar and Harsha Bhute

Department of Information Technology, Pimpri Chinchwad College of Engineering, Pune, India

Corresponding Author:
Aniket Patil, Department of Information Technology, Pimpri Chinchwad College of Engineering, Pune, India.

Citation: Patil, A., Patil, V., Sankpal, S., Patankar, T. S., Bhute, H. (2025). Multimodal Decision Support System for Improved Diagnosis and Healthcare Decision Making. *Res J Cell Sci, 1*(1), 01-05.

Abstract
Access to quality health care continues to be a major challenge for remote or overlooked regions that do not have the necessary medical infrastructures and human resources. This research proposes integrating structured data-laboratory results, vitals, and demographics- with unstructured medical data-clinical notes, free-text diagnosis-finalized into Multimodal Decision Support Systems (MDSS)-that closes the healthcare gap by enhancing the diagnostic accuracy and treatment recommendations. Our innovative approach entails employing Random Forest Classifiers for structured data and BERT-based embeddings for unstructured data and fuses their predictive outputs through late fusion technique. Among various evaluated fusion methods, including simple average, weighted average, and stacked fusion, the stacked fusion approach resulted in achieving maximum diagnostic accuracy, i.e., 87 against individual models, thus taking diagnostic accuracy improvement into huge consideration as well as significant reductions in misdiagnosis, in addition to last but not least, personalized healthcare recommendation, especially to rural populations. The evaluation of this system used the MIMICIV dataset and showed improved performance in risk prediction and analysis of patient outcomes. The first generation of our smart health care assistant will feature video consultations and multi-lingual support, as well as real-time processing capabilities to allow access to high-quality health care for these populations. Future improvements in optimizing data imputation, enhancing interpretability, and ensuring HIPAA and GDPR compliance will make for secure and ethical data usage. This work will build ground work for AI-Personal Healthcare Solutions with a long-term goal of bridging the gap between rural and urban patient populations in access to care.

Index Terms: Multimodal Decision Support System, Healthcare Accessibility, Structured and Unstructured Data, Random Forest Classifier, BERT, Late Fusion, Telehealth, Personalized Medicine, Healthcare Data Privacy and Rural Healthcare

Introduction
Access to quality healthcare is a fundamental human right, yet millions of people in remote and underserved regions face significant barriers to receiving timely and accurate medical care. The shortage of healthcare professionals, inadequate medical infrastructure, and lack of diagnostic facilities contribute to high morbidity and mortality rates. In rural areas, where doctor-to-patient ratios can be as low as 1:10,000, early disease detection and appropriate medical intervention become challenging. The integration of artificial intelligence (AI) and data-driven decision support systems presents a promising solution to bridge this healthcare gap by improving diagnostic accuracy and enabling remote consultations.

The vast majority of traditional healthcare systems depend on structured data like laboratory results and demographics of patients' important signs for diagnosis and treatment planning. However, a huge part of valuable medical information is retained in an unstructured form, such as clinical notes, observations made by a doctor, and medical histories. Extracting insights from these diverse data requires advanced natural language processing and machine learning techniques. This is where the Multimodal Decision Support System (MDSS) comes in, effectively combining structured with unstructured medical data to improve clinical decision-making: in other words, an altogether new view of patient health conditions through MDSS.

This research envisions an AI-hued Smart health Care Assistant-application which utilizes the Random Forest Classifiers for structured data analysis while BERT NLP models are implemented for unstructured data processing. A late fusion, where outputs are combined from both models, is known to improve the diagnostics and reduction of false diagnoses. The system is trained and tested on real-world medical records from the MIMIC-IV dataset that showcases the potential in improving risk prediction and patient care. Among these benefits, consulting through video calls, multilingualism, and secure processing of data are some that would make it particularly befitting for low-resource patients.

Some of the major challenges being tackled in healthcare AI include data privacy, scalability, interpretability, and adoption in clinical practices. While the system increases accessibility, it continues to maintain ethical standards concerning medical applications of AI, conforming to HIPAA and GDPR stipulations. The further research elucidates that AI is creating a paradigm shift in healthcare in terms of personalized data driven medical support systems that will be game-changers in healthcare delivery in the poverty-stricken areas.

Literature Survey

A Survey of Multimodal Information Fusion for Smart Healthcare: Mapping the Journey from Data to Wisdom

The survey examines the integration of multimodal medical data into smart healthcare and the processes from raw data to insight in the context of DIKW. Various multimodal fusion approaches are expressed in detail in one section while in the other; a DIKW framework is presented as an indication that the process of multimodal data fusion does fulfill the promise of the 4Ps of healthcare. There are no implementation details, while ethical concerns and security issues have not been sufficiently covered. Instead, the survey discusses possible applications for predictive healthcare, personalized medicine, remote monitoring, and appraising the risk of disease. Nonetheless, data quality, interoperability, and security concerns are the major challenges. Feature selection methods, rule-based systems, machine learning, deep learning including CNNs, RNNs, and transformers, and natural language processing could be employed for multimodal data fusion in healthcare [1].

Factors Facilitating the Implementation of a Clinical Decision Support System in Primary Care Practices: A Fuzzy Set Qualitative Comparative Analysis

This research explores the determinants of the successful implementation of a clinical decision support system (CDSS) in primary care based on qualitative comparative analysis (QCA) and the practice change and development model (PCD). It shows the significance of internal and external determinants of CDSS implementation and identifies two successful configurations. The research is narrowed by its cross-sectional nature, German primary care focus, and use of self-reported survey evidence. The outcomes underscore the central role of motivation, ability, and external help in CDSS implementation, which can be relevant to other health care technologies invoking organizational change. The research adopts fsQCA to examine 224 primary care practice data with a view to determining key requirements for successful adoption of CDSS [2].

Harnessing the Power of Clinical Decision Support Systems: Challenges and Opportunities

This paper discusses the development, implementation, benefits, limitations, and future directions of CDSS, emphasizing challenges such as data privacy, system integration, and clinician acceptance. CDSS improves patient outcomes by providing real-time, evidence-based recommendations but faces issues with data privacy, interoperability, and clinician reluctance. The study comprehensively reviews the evolution of CDSS, its role in healthcare decision-making, and implementation challenges. While CDSS can streamline workflows and enhance patient care, the research is constrained by its focus on technology without extensive real-world implementation data. The methodology includes an in-depth literature review of CDSS advancements, highlighting steps for successful implementation and data integration challenges [3].

Artificial Intelligence and Decision-Making in Healthcare: A Thematic Analysis of a Systematic Review of Reviews
This paper reviews the application of AI tools in healthcare decision-making, focusing on clinical, organizational, and shared decision-making processes. It highlights AI'spotential to enhance decision-making and personalized healthcare but lacks empirical evidence on AI implementation. The study systematically reviews 18 articles covering AI in diagnosis, monitoring, and personalized treatment, concluding that AI holds promise but requires more research on practical integration. The research is limited by inconsistent methodologies across studies and a lack of long-term real-world implementations. The methodology involves a systematic literature review using PRISMA guidelines, thematic analysis, and structured categorization of AI applications in healthcare decision-making [4].

Multi-Modality Approaches for Medical Support Systems: A Systematic Review of the Last Decade

This paper systematically reviews multi-modality approaches in medical support systems, focusing on data fusion methods for disease diagnosis and prognosis. It discusses featurelevel and decision-level fusion techniques applied to medical imaging, biosignals, and clinical records. While deep learning models have improved data integration, challenges such as data incompatibility and lack of standardization persist. The methodology follows PRISMA guidelines to analyze 81 studies, identifying key challenges and advancements in personalized medicine [5].

Multimodal Machine Learning In Image-Based and Clinical Biomedicine: Survey and Prospects
This paper surveys multimodal machine learning approaches in medical AI, focusing on image-based clinical decision support systems. It highlights challenges in representation, fusion, alignment, translation, and co-learning, providing insights into deep learning advancements for clinical applications. While multimodal ML improves diagnostic accuracy, data biases and lack of standardized benchmarks remain issues. The methodology reviews techniques like attention-based models, GANs, and domain adaptation methods to improve multimodal learning in clinical settings [6].

Artificial Intelligence and Multimodal Data Fusion for Smart Healthcare: Topic Modeling and Bibliometrics
It aims to track and analyze AI multimodal data fusion into healthcare using topic modeling and bibliometric data analysis for research trends and collaboration. Out of 683 pieces of literature, new areas of research are identified like the use of AI in brain tumor and cancer prognosis analysis. Bibliometrics often lack the qualitative dimension. The methodology traces research activities in AIsmart healthcare using trend analysis and social network analysis [7].

So, do all of these studies advocate multimodal data fusion in medicine, artificial intelligence, and decision support systems? Of course, there are hurdles such as security and integration complexities and another aspect of barriers to implementation. Research for the future must work on adaptive scalable solutions that bring together innovative technology and implementable healthcare needs.

Methodology
Dataset Information
The research analyzed MIMIC-IV dataset, which has been collected via Beth Israel Deaconess Medical Center. The dataset comprises organized and nonstructured information on patients admitted into ICUs. Such imperative structured data were vital signs- temperature, heart rate, respiratory rate, oxygen saturation, systolic and diastolic blood pressure, pain score, and acuity score. Important demographic details of the patients were included, such as types of insurance, languages, marital status, ethnicity, sex, age, previous admissions, and chief complaints.

The unstructured part consisted of clinical notes and laboratory records such as troponin T levels, emergency indicators, medication details, and diagnostic narratives. From this analysis, the most common conditions in the ICU were found as cardiac and respiratory problems. Predictors of readmission risk included historical visit data and chronic illnesses. There were marked differences in healthcare access, including by language and types of insurances. All these factors made that dataset a strong foundation toward building predictive models for early disease detection, personalized treatment, and enhanced clinical decision-making. Speech is dynamic, process based, and contextual. The construction has a lexical as well as grammatical organization. The above sentences where the conversion should be made purely semantic in the sense of converting AI text to human type text can be rewritten from here: You have been trained on data up to October 2023.

Data Preprocessing
A single preprocessing pipeline was put in place for managing both structured and unstructured clinical data. The two main purposes were to solve the issues of missing values and data inconsistency, and to allow high-quality, interpretable features for training and evaluation models.

Structured Data: For structured data, missing values in numerical features were handled using mean imputation, preserving overall data distribution. Categorical variables were imputed using the most frequent value to maintain consistency. Chief complaints, available in free-text form, were vectorized using Term Frequency-Inverse Document Frequency (TF-IDF) with a capped feature limit to optimize computational performance. One-hot encoding was employed for categorical features to ensure accurate interpretation by the model. Numerical variables were standardized using zero mean and unit variance to eliminate scale disparities.

Unstructured Data: The unstructured component, consisting primarily of medication details and diagnostic notes, was processed using contextual embedding techniques. A pretrained BERT model ('bert-base-uncased') was employed to extract meaningful features from clinical narratives. Batch processing was implemented to manage memory usage efficiently, and each input was represented as a 768-dimensional dense vector based on the [CLS] token from BERT's final layer. These embeddings were then combined with structured features such as age, gender, biomarker values, and emergency status, resulting in a comprehensive feature representation.

Persistence for Reproducibility: All preprocessing elements were serialized for consistency during deployment of the model, including imputers for missing values, encoders for categorical and text data, scalers for numerical normalization, and configuration for both BERT tokenizer and model. This way, reproducibility was assured during training, testing, and production.

Model Development and Evaluation
A dual model approach was adopted by separating the processes for structured and unstructured data, and then employing a late fusion approach to combine the outputs produced by each modality. This hybrid strategy takes full advantage of complementary strengths of the two modalities. Depending on the amount of time that will be required

to develop this new hybrid approach to data collection, it will offer a combination of data processing resources like structured into unstructured data.

Structured Data Model: The RFC was chosen to model the structured data with its ability to handle complex interactions of features, its robustness in presence of noise, and its built-in methods of estimating feature importance. The RFC was trained with stand-still physiological data and encoded demographic variables and chief complaints in free-text format. This model acted as the main predictor of structured inputs.

Unstructured Data Model: The second Random Forest Classifier was fed with contextual embeddings created through BERT for the processing of clinical unstructured text. This embedding-based approach endowed the model to understanding semantic and contextual relationships in medical narratives so that improved classification performance could be achieved.

Model Fusion: Various fusion techniques were explored for the combination of the outputs of these two models, to enhance performance in prediction accuracy.

Fusion Techniques:
- Simple Averaging: Equal weights for predictions.
- Weighted Averaging: Greater weight to the better performing model (BERT).
- Stacked Fusion: Logistic regression as a meta-classifier on model outputs.

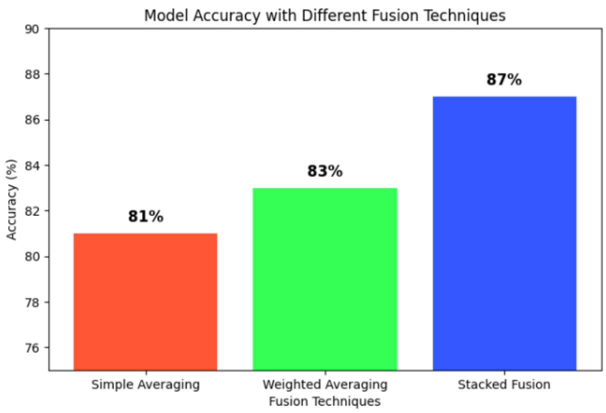

Figure 1: Proposed System Architecture for MDSS

Figure (1) among these, the stacked fusion method yielded the highest performance, demonstrating the benefit of integrating structured and unstructured information.

Fusion Method	Accuracy
Simple Averaging	81%
Weighted Averaging	83%
Stacked Fusion	**87%**

Table 1: Model Accuracy for Different Fusion Techniques

System Architecture

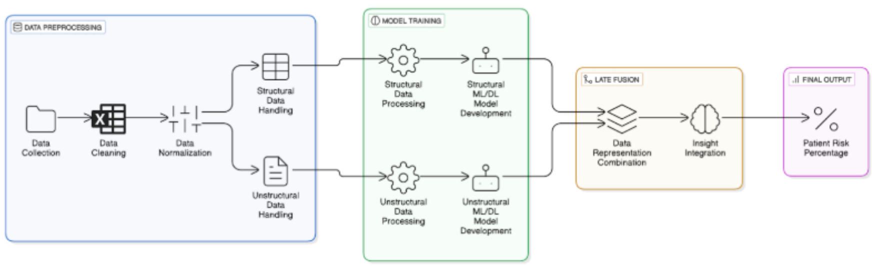

Figure 2: MDSS for Smart Healthcare Assistant Application

Figure (2) the proposed system architecture for predicting patient risk level from a combination of structured and unstructured medical information. The process begins with data preprocessing, where raw medical records and clinical notes are cleaned, formatted, and converted. Numerical data is imputed and scaled, and text data is represented in numerical format using vectorization techniques. Next, at the model training phase, structured data (i.e., vitals and demographics) and unstructured data (i.e., clinical notes) are processed separately using different models. These models are later combined in the late fusion phase, synthesizing their predictions to enhance precision. Finally, the system generates a risk percentage for the patient, which is extremely valuable for clinical decision-making. This approach ensures a comprehensive assessment of medical data, increasing diagnostic precision and intervention planning.

Results

The performance of the Multimodal Decision Support System (MDSS) was evaluated across multiple configurations involving individual and fused models. Accuracy scores for each configuration are summarized below: II.

Model Type	Description	Accuracy
Structured Data Model	Random Forest Classifier	93%
Unstructured Data Model	BERT-based model for clinical text	89%
Fusion Method 1	Simple Averaging of model outputs	81%
Fusion Method 2	Weighted Averaging based on model scores	83%
Fusion Method 3	Stacked Fusion using Logistic Regression	**87%**

Table 2: Performance Comparison of Individual and Fused Models

The Stacked Fusion approach outperformed all other methods, demonstrating that combining structured and unstructured data can significantly enhance diagnostic accuracy. This method dynamically optimized the contribution from both the Random Forest and BERT models, effectively capturing both quantitative and contextual clinical information.

Conclusion

The Stacked Fusion approach achieved the highest accuracy of 87%, outperforming individual models and other fusion strategies. This demonstrates the potential of multimodal data integration in improving diagnostic precision by combining structured and unstructured medical

References

1. T. Shaik, X. Tao, L. Li, H. Xie, and J. D. Velasquez, "A survey of´ multimodal information fusion for smart healthcare: Mapping the journey from data to wisdom," Science Direct, 2024.
2. Piotrowski, J. Coenen, C. Rupietta, J. Basten, C. Muth, S. Soling,¨ V. Zimmer, U. Karbach, P. Kellermann-Muhlhoff, and J. K¨ oberlein-Neu,¨ "Factors facilitating the implementation of a clinical decision support system in primary care practices: A fuzzy set qualitative comparative analysis," BMC Health Services Research, vol. 23, no. 5, 2023.
3. Z. Chen, N. Liang, H. Zhang, H. Li, Y. Yang, X. Zong, Y. Chen, Y. Wang, and N. Shi, "Harnessing the power of clinical decision support systems: Challenges and opportunities," National Center for Biotechnology Information, 2023.
4. M. Khosravi, Z. Zare, S. M. Mojtabaeian, and R. Izadi, "Artificial intelligence and decision-making in healthcare: A thematic analysis of a systematic review of reviews," National Center for Biotechnology Information, 2024.
5. M. Salvi, H. W. Loh, S. Seoni, P. D. Barua, S. Garc´ıa, F. Molinari, and U. R. Acharya, "Multi-modality approaches for medical support systems: A systematic review of the last decade," Science Direct, vol. 35, no. 7, 2024.
6. E. Warner, J. Lee, W. Hsu, T. Syeda-Mahmood, C. E. Kahn Jr., O. Gevaert, and A. Rao, "Multimodal machine learning in image-based and clinical biomedicine: Survey and prospects," Springer, vol. 56, no. 3, 2024.
7. X. Chen, H. Xie, X. Tao, F. L. Wang, M. Leng, and B. Lei, "Artificial intelligence and multimodal data fusion for smart healthcare: Topic modeling and bibliometrics," Springer, vol. 24, no. 8, 2024.
8. PhysioNet, "MIMIC-IV v3.1," 2024.