MediGuide: AI Driven Personalized Healthcare And Diagnostics System

1st Siddhesh Hagawane
Department of Artificial Intelligence and Data Science
New Horizon Institute of Technology & Management
Thane, India

2nd Aster Noronha
Department of Artificial Intelligence and Data Science
New Horizon Institute of Technology & Management
Thane, India

3rd Adarshkumar Gautam
Department of Artificial Intelligence and Data Science
New Horizon Institute of Technology & Management
Thane, India

4th Dr. Megha Gupta
Department of Artificial Intelligence and Data Science
New Horizon Institute of Technology & Management
Thane, India

***Abstract*—The integration of AI in healthcare has revolutionized disease prediction and personalized recommendations. Mediguide, an AI-driven system, predicts diseases using machine learning (ML) and deep learning (DL) algorithms, analyzing 132 symptoms to diagnose 42 diseases. It also uses image-based models for identifying 8 skin conditions and provides dietary, medication, and lifestyle advice. A chatbot interface enhances accessibility, enabling personalized interactions. Mediguide addresses the growing demand for quick medical guidance amid a shortage of healthcare professionals. However, challenges persist in improving adaptive recommendations and prediction confidence. This scalable approach has potential but requires further refinement for better personalization and reliability.**

***Index Terms*—Knowledge Graph, Retrieval-Augmented Generation, Large Language Models, Vector Search, Graph Search, Information Extraction, Cross-Document Context, Structured Representation.**

I. Introduction

Healthcare AI solutions have gained traction in recent years, offering efficient and scalable approaches to disease diagnosis and patient care. Traditional diagnostic methods often require significant medical expertise and time, limiting accessibility in remote areas. Mediguide aims to bridge this gap by providing an AI-powered disease prediction system that can analyze symptoms and suggest relevant treatments. The system consists of three key components: 1.Disease Prediction Model: Uses symptom-based analysis and deep learning for image-based disease detection. 2.Chatbot: Enhances user interaction and provides a personalized experience. 3.Recommendation System: Suggests medications, dietary plans, and workout routines.

Mediguide leverages large datasets of medical conditions and symptom correlations to generate predictions. The disease prediction model operates on structured medical databases, mapping symptom clusters to likely diagnoses. By integrating AI-based models, Mediguide can handle vast amounts of data efficiently and produce results in real-time, making it a valuable tool for users seeking quick medical guidance. Despite its potential, challenges such as personalization, adaptive recommendations, and confidence scoring need to be addressed. This paper explores Mediguide's development, methodology, and key results, highlighting future improvements for AI-driven medical assistance.

II. Related Work

In recent years, numerous studies have explored the integration of machine learning and deep learning for disease prediction and healthcare recommendation systems. Researchers have utilized various algorithms to enhance accuracy, efficiency, and adaptability in medical decision-making.

Kurian et al. [1] conducted a comparative analysis of several machine learning models, including Support Vector Machine (SVM), Random Forest, Decision Tree, and XGBoost, to determine their effectiveness in disease prediction and medicine recommendation. Their study emphasizes the potential of these models in improving healthcare decision-making. Similarly, Gomathy et al. [2] provided a systematic review of disease prediction techniques, highlighting the strengths and limitations of Na¨ıve Bayes, Decision Tree, and SVM algorithms in healthcare applications.

Deep learning techniques have further revolutionized medical diagnostics by improving prediction accuracy. Gaurav et al. [3] proposed a neural network-based approach using Long Short-Term Memory (LSTM), Random Forest, and SVM, achieving notable improvements in predicting health conditions. Another study by Grampurohit et al. [4] introduced the ID3 Decision Tree Algorithm, which prioritizes frequently occurring symptoms to enhance diagnostic precision. Additionally, Park et al. [5] employed ensemble learning models such as LightGBM and XGBoost to process laboratory test data for disease prediction, demonstrating high accuracy levels.

Medical recommendation systems have also gained significant traction in healthcare research. Bhidve et al. [6] implemented a neural network-based model combined with Random Forest, allowing for efficient medicine recommendations based on diagnosed conditions. Similarly, Mohan et al. [7] proposed a CRNN-based approach with self-attention mechanisms and incorporated sentiment analysis of medicine reviews to refine medication recommendations and enhance patient satisfaction.

In addition to disease prediction and recommendation systems, AI-powered chatbots are increasingly used for healthcare assistance. Gadge et al. [8] developed a deep learning-driven chatbot using Natural Language Toolkit (NLTK) and Bag of Words (BoW) to provide real-time medical guidance and consultations. Their work highlights the growing role of AI-driven conversational agents in supporting patient care and addressing general health inquiries.

Recent advancements in AI have led to the development of sophisticated healthcare assistants. For instance, Microsoft's Dragon Copilot is an AI assistant designed to alleviate clinicians' administrative burdens by automating documentation tasks and providing reliable medical information [9]. Similarly, Movano introduced EvieAI, an AI chatbot trained exclusively on peer-reviewed medical journals to offer accurate health and wellness information, ensuring high accuracy and adherence to medical best practices [10].

Building upon these advancements, Mediguide integrates XGBoost for symptom-based disease prediction, MobileNetV2 for image-based skin disease detection, and Llama 3.2 for chatbot-driven healthcare assistance. Unlike previous studies, Mediguide adopts a modular architecture where disease prediction, recommendation, and chatbot functionalities operate independently, offering enhanced flexibility and accuracy in healthcare support. The recommendation system follows a rule-based approach, providing general medication and dietary suggestions based on the diagnosed disease. Users seeking personalized assistance can leverage the chatbot, which utilizes natural language processing (NLP) to interact dynamically with users. Future enhancements aim to refine chatbot personalization, improve disease coverage, and integrate advanced AI techniques for a more comprehensive healthcare system.

III. Methodology

Mediguide is a healthcare recommendation system that integrates machine learning, deep learning, and natural language processing (NLP) to provide disease predictions and general recommendations. The system consists of three standalone modules: a disease prediction model, a chatbot interface, and a recommendation system. The disease prediction model utilizes XGBoost for symptom-based disease classification, trained on a dataset comprising 132 symptoms mapped to 42 medical conditions. Additionally, a deep learning-based Convolutional Neural Network (CNN) using MobileNetV2 is employed for image-based skin disease detection, leveraging transfer learning from ImageNet. Mediguide is a human-inspired AI-driven healthcare recommendation system designed to mimic real-world medical decision-making. The conceptual framework of Mediguide is illustrated in the diagram, which draws parallels

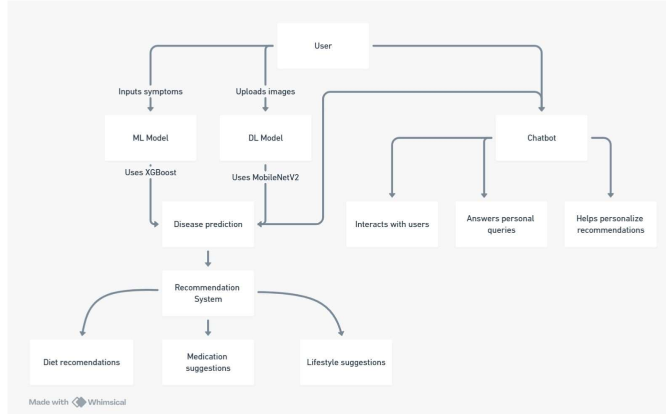

Fig. 1. Architecture of System

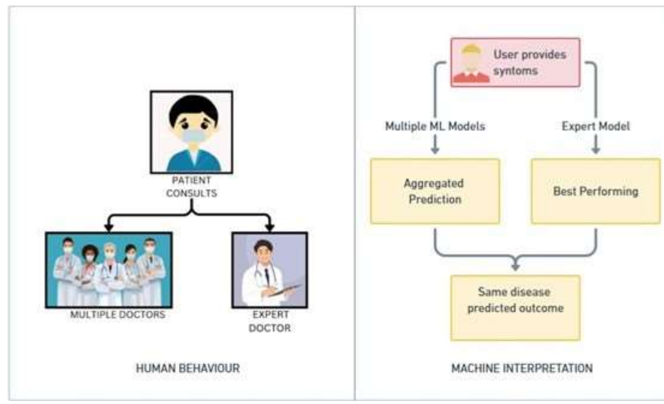

Fig. 2. Diagnosis Framework

between human medical consultation behavior and machine-based diagnosis. In traditional healthcare, a patient may consult multiple doctors to reach a diagnosis, and if one treatment fails, they seek a second opinion. Our approach initially mirrored this process by training multiple machine learning models, including Random Forest, XGBoost, Support Vector Machine, and Multilayer Perceptron. The predictions from these models were aggregated to determine the most agreed-upon disease as the final diagnosis. However, through extensive experimentation, XGBoost consistently outperformed the other models, aligning closely with the multi-model consensus approach while offering superior accuracy and reliability. As a result, the system was optimized to rely primarily on XGBoost for symptom-based disease prediction.

Refering to Figure 2 As depicted in the diagram, the left side represents human behavior, where a patient seeks consultations from multiple doctors or an expert doctor. The right side illustrates Mediguide's approach, where a user provides symptoms and multiple machine learning models generate predictions. These predictions are either aggregated or validated against the best-performing model, resulting in a final disease prediction. This structured methodology ensures robustness and reliability, akin to human medical decision-making. The chatbot acts as a healthcare assistant, providing

TABLE I
XGBoost Model Performance

Metric	Precision	Recall	F1-score	Samples
Accuracy	-	-	0.95	993
Macro Avg	0.96	0.95	0.94	993
Weighted Avg	0.96	0.95	0.94	993

TABLE II
MobileNetV2 Model Performance

Metric	Precision	Recall	F1-score	Samples
Accuracy	-	-	0.95	233
Macro Avg	0.95	0.95	0.95	233
Weighted Avg	0.95	0.95	0.95	233

general medical guidance and responding to user queries. It is powered by Llama 3.2, ensuring interactive, context-aware assistance. The chatbot operates independently and is not directly integrated with the prediction or recommendation systems. The recommendation system follows a rule-based approach, offering medication, diet, and lifestyle recommendations strictly based on the diagnosed disease. For users seeking personalized recommendations, the chatbot serves as an alternative.

A. Dataset

The dataset used in Mediguide was curated from medical literature and expert consultations. The symptom-based disease prediction model was trained using the SymbiPredict Dataset, a structured collection linking symptoms to various diseases, inspired by methodologies from institutions such as the Centers for Disease Control and Prevention (CDC). For skin disease classification, the Skin Disease Dataset from Kaggle was utilized, containing labeled images of multiple dermatological conditions. These datasets provided a diverse and reliable foundation for model training and evaluation. The system was developed using Python and deployed with Streamlit for an interactive user interface. Since each module also operates independently, they offer flexibility in usage. Model evaluation was conducted using accuracy, precision, recall, and F1-score for disease prediction, while chatbot relevance and recommendation system effectiveness were assessed via user feedback and expert validation. Ethical considerations were also prioritized, ensuring compliance with healthcare data privacy standards and responsible AI practices. Future improvements aim to enhance chatbot personalization and expand disease coverage for a more comprehensive healthcare experience.

IV. Results

The proposed disease prediction model was evaluated using multiple performance metrics, including accuracy, precision, recall, and F1-score. Additionally, visual analysis through confusion matrices and precision-recall curves was conducted to gain deeper insights into model performance.

A. XGBoost Model Performance

Table I represents the performance of the XGBoost-based disease prediction model, which achieved a high classification accuracy of 95 percentage, demonstrating its effectiveness in predicting 42 diseases based on input symptoms. The macro and weighted averages of precision, recall, and F1-score further validate its robustness

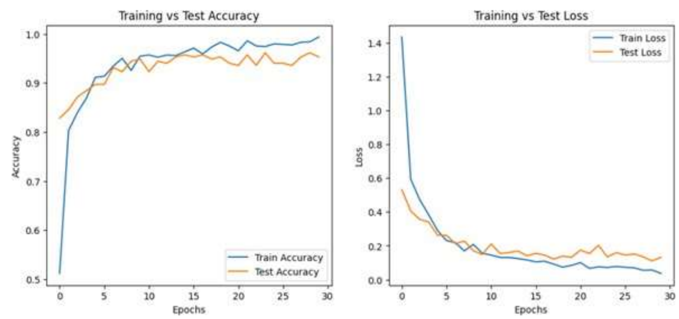

Fig. 3. Training vs.Test Accuracy Loss Graphs for Model Stability

• Macro Average: This metric treats all disease classes equally, averaging precision, recall, and F1-score across all categories.

• Weighted Average: Adjusts for class imbalance by assigning higher weights to frequent disease classes, ensuring a fair representation in overall performance.

B. MobileNetV2 Model Performance

The Table II represents the deep learning-based image classification model for skin disease detection, utilizing MobileNetV2, was evaluated using accuracy, precision, recall, and F1-score. The model demonstrated a high testing accuracy of 95.28 percent, indicating strong predictive capabilities for skin disease classification.

The training and testing accuracy trends over 30 epochs indicate effective learning, with training accuracy reaching 99.24 percentage and testing accuracy stabilizing at 95.28 percentage. The loss curves suggest minimal overfitting, confirming the model's generalization capability.

C. Feature Importance Analysis

The Figure 4 helps to understand the most significant symptoms in disease classification, a feature importance analysis was conducted using the XGBoost model. The top 20 most relevant symptoms influencing predictions were identified and visualized.

D. Error Analysis and Future Improvements

While the model achieves high performance, certain diseases with overlapping symptoms, such as Dengue and Malaria, or Eczema and Psoriasis, pose classification challenges. Future work will incorporate:

• Enhanced feature extraction using deep learning for better symptom representation.

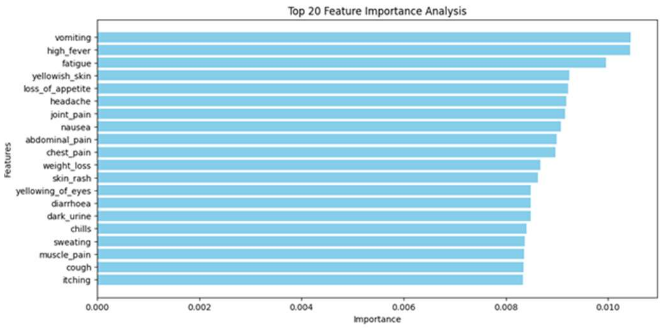

Fig. 4. Top 20 Critical Symptoms for Disease Prediction

• Integration of medical history and lab test results for improved context awareness.

• Application of explainable AI (XAI) to increase interpretability for healthcare professionals.

The model provides a reliable and high-accuracy disease prediction system, offering potential applications in telemedicine and automated diagnosis tools. Future enhancements will focus on reducing misclassification rates and increasing interpretability for real-world deployment.

V. Conclusion

Mediguide is a powerful AI-driven healthcare solution designed to predict diseases by analyzing symptoms and medical images. With the capability to diagnose 42 diseases, it provides users with quick and automated medical insights, reducing the time required for initial diagnosis. By leveraging machine learning, Mediguide identifies patterns in patient data and recommends possible treatments based on its analysis. This approach enhances accessibility to healthcare, especially in remote or underserved areas where medical professionals may not be readily available. However, despite its efficiency, the system currently lacks personalized recommendations, which limits its ability to tailor advice based on an individual's medical history, lifestyle, or genetic factors. Additionally, the absence of confidence scores in its predictions makes it difficult for users to assess the reliability of the diagnosis, potentially leading to uncertainty in decision-making.

To improve its effectiveness, future enhancements will focus on adaptive learning, allowing Mediguide to refine its predictions based on user feedback and real-world medical data. Deeper integration of a chatbot interface will enhance user interaction, enabling more detailed symptom analysis and better guidance. Additionally, advancements in image analysis will improve the accuracy of disease detection, ensuring more precise diagnoses. By incorporating these improvements, Mediguide aims to offer a more intelligent and responsive healthcare solution that provides reliable medical assistance with greater confidence. As AI continues to transform the healthcare industry, solutions like Mediguide highlight the potential of technology in making healthcare more accessible, efficient, and personalized, ultimately leading to better patient outcomes.

References

[1] L. Kurian, E. K. J., and G. Joseph K., ”Disease prediction and medicine recommendation systems: A comparative analysis on learning algorithms.”

[2] C. K. Gomathy and M. A. Rohit Naidu, ”The prediction of disease using machine learning (IJSREM).”

[3] K. Gaurav, A. Kumar, P. Singh, A. Kumari, and T. Suryawanshi, ”Human disease prediction using machine learning techniques and real-life parameters.”

[4] S. Grampurohit and C. Sagarnal, ”Disease prediction using machine learning algorithms.”

[5] D. J. Park, M. W. Park, H. Lee, Y. J. Kim, Y. Kim, and Y. H. Park, ”Development of machine learning model for diagnostic disease prediction based on laboratory test.”

[6] S. Bhidve, J. Lamkhade, and A. Mavale, ”Medicine recommendation system.”

[7] G. B. Mohan, R. P. Kumar, and E. R., ”Medical recommendations: Leveraging CRNN with self-attention mechanism for enhanced systems.”

[8] H. Gadge, V. Tode, S. Madane, P. Kachare, and A. Deokar, ”A Chatbot for medical purposes using deep learning.”

[9] ”Microsoft's new Dragon Copilot is an AI assistant for healthcare,” The Verge, 2025. (theverge.com)

[10] ”This smart ring is rolling out an AI chatbot trained on medical journals,” The Verge, 2025. (theverge.com)