ISSN: 2582-7219

International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Impact Factor: 8.206

Volume 8, Issue 10, October 2025

© 2025 IJMRSET | Volume 8, Issue 10, October 2025| DOI:10.15680/IJMRSET.2025.0810039

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

AI-Based Medical Diagnosis System

Prof. Dr. Naresh Thoutam[1], Rambabu Yadav[2], Siddhant Mahesh Pawar[3], Suhani Sagarsingh Pardeshi[4], Sakshi Pravin Ugale[5]

Department of Computer Engineering, Sandip Institute of Technology and Research Centre, Nashik, India[1-5]

ABSTRACT: The integration of Artificial Intelligence (AI) into healthcare has significantly transformed the medical diagnosis landscape by enabling faster, more reliable, and accessible decision-support tools. This project focuses on the development of an AI-based Medical Diagnosis System that leverages both Machine Learning (ML) and Deep Learning (DL) models to predict various diseases, including diabetes, heart disease, kidney disease, liver disease, breast cancer, malaria, and pneumonia. The system processes both structured medical data and medical images, providing patients with real-time predictions and facilitating communication with doctors through appointment booking and email support. Random Forest classifiers are employed for structured datasets, while Convolutional Neural Networks (CNNs) are applied to medical imaging tasks. The system aims to reduce diagnostic delays, assist healthcare professionals in decision-making, and improve patient outcomes by offering an easy-to-use web interface. Although not a replacement for medical experts, the system serves as an initial screening tool that enhances healthcare accessibility, especially in remote and underserved areas. The project demonstrates the potential of AI in medical diagnostics and establishes a scalable framework for integrating additional diseases in the future.

KEYWORDS: Artificial Intelligence, Machine Learning, Deep Learning, Random Forest, CNN, Medical Diagnosis, Healthcare

I. INTRODUCTION

The healthcare sector has witnessed a paradigm shift with the adoption of artificial intelligence technologies for disease prediction and diagnosis. Traditional medical diagnosis often involves time-consuming laboratory tests, manual analysis, and significant dependence on medical professionals, which can result in delays and misdiagnosis. AI-based systems provide an effective alternative by utilizing algorithms trained on large datasets to deliver accurate and faster results.

Medical diagnosis systems powered by AI are designed to process diverse forms of input such as clinical parameters, imaging data, and patient history. These systems employ machine learning classifiers to handle numerical datasets and deep learning models for image-based predictions. The integration of multiple diseases within a single system provides a unified platform for patients, reducing the need for separate diagnostic tools for different conditions.

Furthermore, AI-based medical systems not only provide diagnostic predictions but also facilitate healthcare accessibility through additional features such as doctor appointment booking and patient-doctor communication. This makes the system more practical for real-world healthcare applications, where convenience and efficiency are critical.

The proposed system is a step towards bridging the gap between technology and healthcare by providing decision-support tools that can assist doctors in early disease detection while ensuring patients receive timely medical intervention.

II. LITERATURE SURVEY

Recent advancements in AI have contributed to significant improvements in medical diagnosis. Chen et al. (2024) proposed an AI-driven multimodal system integrating structured and unstructured data for diagnostic accuracy. Gupta and Mehra (2024) applied deep learning models for pneumonia detection using chest X-ray images, achieving higher precision compared to traditional methods. Patel and Shah (2023) developed a hybrid Random Forest and CNN framework for kidney and liver disease prediction. Zhang et al. (2023) introduced a hybrid system for diabetes and heart disease detection, enhancing predictive reliability. Kumar and Singh (2023) worked on breast cancer prediction models using AI, while Hussain and Ali (2023) applied CNNs to malaria parasite detection in blood smear images.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Thomas and George (2022) presented a broad survey of AI applications in healthcare, emphasizing challenges in data security and ethics. Li and Zhou (2022) applied ensemble learning to cardiovascular disease prediction. Prasad and Nair (2022) focused on chronic kidney disease classification using ML. Singh and Arora (2022) proposed CNN-based pneumonia detection. Zhang and Wu (2021) highlighted opportunities and challenges in AI healthcare systems. Sharma and Saini (2021) applied ML techniques for early diabetes detection. Khan and Rehman (2021) focused on CNNs for breast cancer classification. Gupta and Yadav (2020) used hybrid models for heart disease prediction.

Alam and Akhtar (2020) explored predictive analytics in healthcare, while Rajkomar et al. (2019) discussed ML in medicine. Esteva et al. (2019) outlined the role of deep learning in healthcare. Miotto et al. (2018) highlighted challenges and opportunities for AI in healthcare. Choi et al. (2018) and Deo (2015) further emphasized the promise of machine learning in transforming diagnosis and decision-making in medicine.

Problem Statement

Healthcare systems often face delays in diagnosis due to manual testing, limited availability of skilled professionals, and the reliance on separate diagnostic tools for different diseases. This increases the risk of late detection, misdiagnosis, and higher treatment costs. Patients in rural or underserved areas are particularly disadvantaged due to the lack of access to advanced diagnostic facilities. The problem lies in creating a unified, intelligent, and accessible system that can analyze both structured patient data and medical images to predict multiple diseases accurately. Such a system should not only support medical professionals in decision-making but also provide patients with easy access to diagnosis and communication with doctors. Therefore, the core challenge is to develop an AI-based medical diagnosis system that combines machine learning and deep learning models, integrates user-friendly features, ensures scalability for additional diseases, and maintains accuracy while being practical for real-world healthcare applications.

III. METHODOLOGY

The methodology of the proposed system involves multiple stages, beginning with **data collection and preprocessing**. Structured datasets for diseases like diabetes, heart disease, liver disease, kidney disease, and breast cancer are collected from medical repositories, while medical image datasets for malaria and pneumonia are obtained from imaging datasets. Preprocessing techniques such as normalization, missing value handling, and image resizing are applied to ensure uniformity.

Next, **model training** is carried out using machine learning algorithms such as Random Forest for structured data and CNNs for image data. Random Forest is chosen for its robustness and ability to handle medical datasets with multiple attributes. CNNs are applied for image classification tasks, leveraging their strength in feature extraction and pattern recognition.

The **system development phase** integrates these models into a Flask-based web application. Patients can input their medical data or upload images through the web interface, which is processed by the backend to generate predictions. The system also includes modules for booking appointments and sending email notifications, enhancing practical usability.

Finally, the **evaluation phase** involves testing the system on benchmark datasets and validating results with metrics such as accuracy, recall, and F1-score. Continuous improvements are planned through hyperparameter tuning, cross-validation, and the addition of more diseases to enhance system scalability.

Proposed System

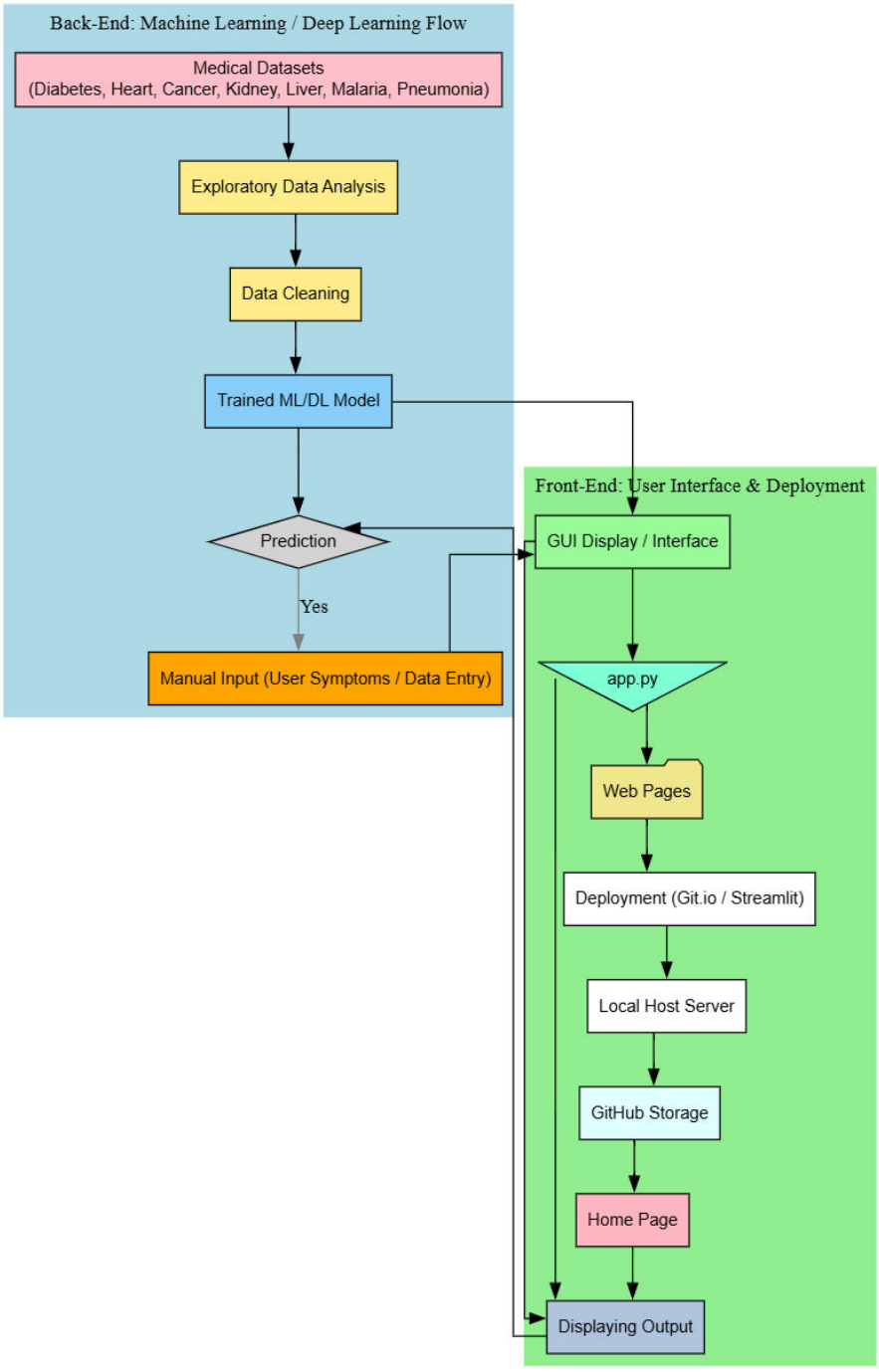
The proposed system is a web-based platform powered by machine learning and deep learning models for medical diagnosis. Patients can input structured health data or upload medical images, which are processed by the backend models. Random Forest is employed for structured data predictions, while CNNs are used for image classification tasks. The results are displayed in real time, and patients are given the option to book appointments or connect with doctors via email. The system architecture consists of three layers: data input, model processing, and result output. This modular structure allows scalability for future enhancements.

© 2025 IJMRSET | Volume 8, Issue 10, October 2025| DOI:10.15680/IJMRSET.2025.0810039

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Objectives

- To develop an AI-based medical diagnosis system capable of predicting multiple diseases.
- To integrate machine learning and deep learning models for handling both structured and unstructured data.
- To provide a user-friendly interface that enables patients to input data, upload images, and view results in real time.
- To enhance healthcare accessibility by including appointment booking and communication features.
- To create a scalable system framework that can be expanded with additional diseases in the future.
- To evaluate system accuracy and reliability using medical datasets and standard performance metrics.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

IV. SOFTWARE AND HARDWARE REQUIREMENTS

Software Requirements:

- Programming Language: Python 3.10
- Frameworks: Flask, TensorFlow/Keras, Scikit-learn
- Libraries: Pandas, NumPy, OpenCV, Matplotlib
- Database: SQLite/MySQL
- Development Environment: VS Code, Jupyter Notebook

Hardware Requirements:

- Processor: Intel i5/i7 or equivalent
- RAM: Minimum 8 GB (16 GB recommended for DL tasks)
- GPU: NVIDIA GPU with CUDA support (for CNN training)
- Storage: 500 GB HDD / 256 GB SSD
- Internet Connectivity: Required for dataset access and model updates

V. CONCLUSION

The AI-based Medical Diagnosis System provides a unified, intelligent, and accessible approach to disease prediction by integrating Random Forest classifiers for structured data and CNNs for image analysis. The system supports early diagnosis of multiple diseases, assists medical professionals in decision-making, and enhances patient accessibility through a user-friendly web interface. While not a replacement for medical expertise, it serves as a reliable decision-support tool that reduces diagnostic delays and improves healthcare outcomes. This project demonstrates the transformative role of AI in medical applications and sets the foundation for future intelligent healthcare systems.

VI. FUTURE SCOPE

- Addition of more diseases such as Alzheimer's, Parkinson's, and skin cancer.
- Integration with IoT and wearable devices for real-time health monitoring.
- Incorporation of Electronic Health Records (EHR) for personalized predictions.
- Use of Natural Language Processing (NLP) for symptom-based diagnosis.
- Deployment on cloud platforms for scalability and global accessibility.
- Enhancing privacy and security through federated learning techniques.

REFERENCES

[1] J. Chen, L. Wang, and J. Xu, "AI-driven diagnostic systems for multimodal healthcare data integration," *IEEE Transactions on Artificial Intelligence*, vol. 5, no. 2, pp. 145–158, 2024.
[2] A. Gupta and R. Mehra, "Deep learning-based pneumonia detection using chest X-ray images," *Biomedical Signal Processing and Control*, vol. 92, p. 105438, 2024.
[3] K. Patel and M. Shah, "Random forest and deep CNN-based hybrid system for liver and kidney disease prediction," *Computers in Biology and Medicine*, vol. 163, p. 107153, 2023.
[4] Y. Zhang, X. Li, and H. Chen, "A hybrid machine learning framework for diabetes and heart disease diagnosis," *Expert Systems with Applications*, vol. 229, p. 120441, 2023.
[5] P. Kumar and S. Singh, "AI-powered breast cancer detection using machine learning and deep learning techniques," *International Journal of Medical Informatics*, vol. 176, p. 105050, 2023.
[6] M. Hussain and T. Ali, "Deep neural networks for malaria parasite detection from blood smear images," *Computers in Medicine and Biology*, vol. 25, no. 4, pp. 275–288, 2023.
[7] J. Thomas and P. George, "Medical diagnosis using artificial intelligence: A survey of models and applications," *Journal of Healthcare Engineering*, vol. 2022, pp. 1–15, 2022.
[8] W. Li and X. Zhou, "Hybrid ensemble learning methods for cardiovascular disease prediction," *Applied Intelligence*, vol. 52, no. 11, pp. 12453–12465, 2022.
[9] R. Prasad and A. Nair, "Predicting chronic kidney disease using machine learning classifiers," *Procedia Computer Science*, vol. 200, pp. 662–671, 2022.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

[10] R. Singh and N. Arora, "A CNN-based approach for pneumonia detection from medical imaging," *IEEE Access*, vol. 10, pp. 12658–12668, 2022.

[11] Y. Zhang and L. Wu, "A survey on AI in healthcare: Opportunities and challenges," *ACM Computing Surveys*, vol. 54, no. 5, pp. 1–36, 2021.

[12] M. Sharma and R. Saini, "Machine learning models for early detection of diabetes mellitus," *Journal of Biomedical Informatics*, vol. 118, p. 103791, 2021.

[13] A. Khan and S. Rehman, "Deep convolutional neural networks for breast cancer image classification," *Pattern Recognition Letters*, vol. 146, pp. 27–34, 2021.

[14] S. Gupta and P. Yadav, "Hybrid machine learning approaches for heart disease prediction," *International Journal of Computer Applications*, vol. 975, no. 8887, pp. 45–52, 2020.

[15] M. Alam and M. Akhtar, "AI in healthcare: Predictive analytics for disease diagnosis," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 8, pp. 123–130, 2020.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com