IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 6, October 2026
Impact Factor: 8.2

Medical Diagnosis System Using Artificial Intelligence (AI)

Shinde Sarangi Ashok[1], Prof. A. B. Jadhav[2] Prof. D. K. Gund[3] and Dr. H. B. Jadhav[4]
Student, Department of Computer Engineering[1]
Professor (Guide), Department of Computer Engineering[2]
Professor (Co Guide), Department of Computer Engineering[3]
HOD, Department of Computer Engineering[4]
Adsul's Technical Campus, Chas, Ahilyanagar, India

Abstract: *Artificial Intelligence (AI) is increasingly being explored in healthcare because it can process clinical information, detect meaningful trends, and offer additional evidence for diagnostic decisions. This study reviews an AI-supported medical diagnosis framework and considers methods such as Machine Learning (ML), Deep Learning (DL), and Natural Language Processing (NLP). It discusses algorithms such as Logistic Regression, Decision Trees, Random Forest, Support Vector Machines, and Artificial Neural Networks. The workflow includes gathering and cleaning patient data, identifying useful features, training a model, generating predictions, and measuring performance through accuracy, precision, sensitivity, specificity, and the F1-score. It also examines challenges related to data privacy, bias, interpretability, and clinical validation. Based on the reviewed literature, AI tools may assist clinicians when developed and used responsibly; final medical decisions must remain under qualified professional supervision*

Keywords: AI in Healthcare, Computer-Aided Diagnosis, Machine Learning, Deep Neural Networks, Clinical Support Tools, Medical Image Processing, Health Risk Estimation

I. INTRODUCTION

Medical diagnosis is an essential part of healthcare. It involves identifying a disease or medical condition by analysing a patient's symptoms, medical history, physical examination, laboratory reports, and diagnostic images. A reliable identification of a condition allows clinicians to choose suitable care and can contribute to better patient outcomes. In conventional practice, physicians consider a patient's examination findings alongside several sources of medical information before reaching a diagnostic conclusion. While this clinical process remains essential, reviewing extensive patient information can demand significant time and specialised knowledge. When a condition is recognised late, appropriate care may also be delayed.

AI has gained attention as a potential tool for assisting clinicians with diagnostic tasks. AI systems can analyse medical data, recognise patterns, and generate predictions based on previously collected information. Machine Learning methods identify relationships from training examples, whereas Deep Learning models can learn detailed representations from images, signals, and other data types.AI-based diagnostic systems have several potential applications, including the interpretation of radiographs, MRI and CT images, laboratory findings, and electronic patient records. These systems can help identify suspicious patterns and prioritise cases for further clinical examination. However, AI-generated predictions are not always correct. The reliability of these tools is influenced by dataset quality, model design, training procedures, and how closely the deployment population matches the data used for development. Therefore, clinical validation and human supervision are necessary before an AI system is used in real healthcare settings.

The main purpose of this research paper is to explain the working principles, algorithms, applications, advantages, limitations, and future scope of AI-based medical diagnosis systems.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 6, October 2026
Impact Factor: 8.2

Objectives of the Study

The main objectives of this study are:

- To understand the role of Artificial Intelligence in medical diagnosis.
- To study Machine Learning and Deep Learning algorithms used in healthcare.
- To outline a possible architecture for a medical diagnostic-support tool powered by AI.
- To describe how data are prepared, models are trained, and diagnostic estimates are produced.
- To identify suitable performance evaluation metrics.
- To discuss the challenges and limitations of AI in healthcare.
- To explore future opportunities for AI-based diagnostic technologies.

Scope of the Study

This study focuses on the use of Artificial Intelligence for clinical decision support, medical-image analysis, disease-risk prediction, and healthcare-data processing.

The framework is designed to examine relevant patient information and present estimated results that healthcare professionals can assess. It is not designed to replace doctors or independently determine treatment.

II. EXISTING SYSTEM

Traditional medical diagnosis depends on the knowledge and experience of healthcare professionals. Doctors collect information about symptoms, examine patients, review laboratory reports, and interpret medical images before reaching a conclusion.

This process is essential, but it can become difficult when healthcare professionals must analyse large amounts of information within a limited time.

Computer-based healthcare systems have been developed to improve the organisation and analysis of medical information. These systems include electronic health records, rule-based diagnostic tools, laboratory information systems, and clinical decision-support systems.

Rule-based systems use predefined conditions to generate recommendations. For instance, the software could flag a laboratory measurement when it falls outside a defined reference limit. Such systems can be useful, but their effectiveness depends on the rules provided by experts.

Machine Learning introduces another approach by learning patterns from labelled examples. Instead of relying entirely on manually written rules, a trained model can estimate the likelihood of a particular outcome based on input data.

A. Limitations of the Existing System

The traditional and rule-based approaches have several limitations:

Manual interpretation can be time-consuming.

Large medical datasets are difficult to analyse manually.

Rule-based systems may not handle complex relationships between multiple variables.

Missing or incorrect patient information can affect decisions.

Healthcare organisations may store records in varying formats and follow different operational processes.

Some diagnostic conditions require specialised knowledge and further investigations.

B. Need for an AI-Based System

An AI-based system can help analyse large datasets, identify patterns, and provide additional information to healthcare professionals. It may help rank cases for review and draw clinicians' attention to findings that warrant closer examination.

However, the system must be tested carefully. Any model-generated estimate should be interpreted together with the person's symptoms, clinical examination, previous health records, and all other relevant evidence.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 6, October 2026
Impact Factor: 8.2

III. PROPOSED SYSTEM

This proposed AI-assisted diagnostic framework analyses permitted patient records and produces decision-support information for assessment by an appropriately qualified clinician

The system combines data pre-processing, feature extraction, Machine Learning algorithms, prediction, and performance evaluation into a single workflow.

The proposed system follows these major stages:

Data collection.

Data validation and pre-processing.

Feature extraction.

AI model selection and training.

Disease-risk prediction.

Result presentation.

Clinical review and monitoring.

A. System Architecture

The system architecture consists of the following modules.

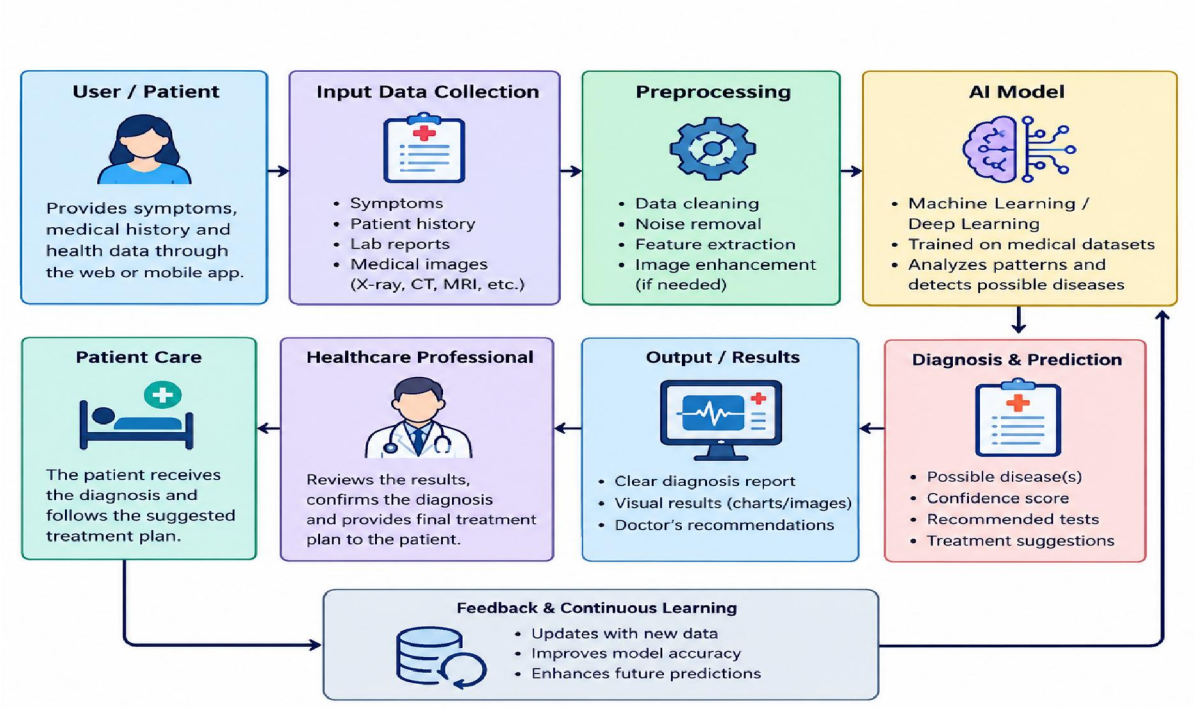


Fig. 1. System Architecture of the Proposed AI-Based Medical Diagnosis System.

B. Working of the Proposed System

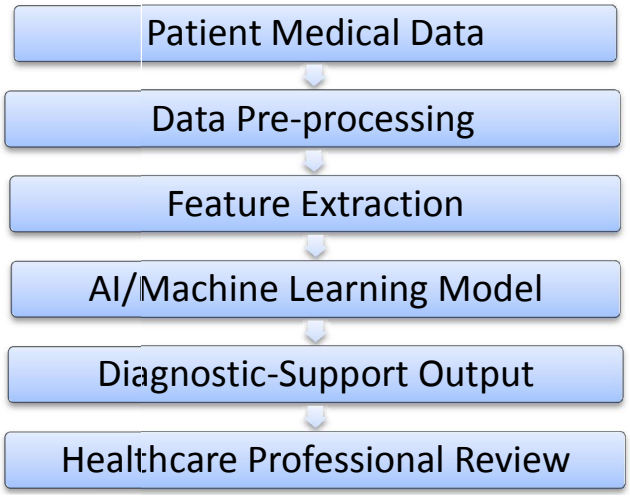
Step 1: Data Collection

The system receives relevant information, such as symptoms, medical history, laboratory measurements, or medical images, from approved sources.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 6, October 2026
Impact Factor: 8.2

Step 2: Data Pre-processing
The collected data is checked for missing values, duplicate records, inconsistent formats, and invalid measurements. Appropriate pre-processing methods are applied according to the data type.
Step 3: Feature Extraction
Important information is represented numerically or transformed into features that the selected AI algorithm can process.
Step 4: Model Training
A suitable Machine Learning or Deep Learning model is trained using labelled data. The training process helps the model learn patterns associated with the target condition.
Step 5: Prediction
The trained model analyses new input data and produces an estimated class, probability, or risk score.
Step 6: Result Display
The system presents its output with relevant warnings and limitations. It should not describe an unverified prediction as a confirmed diagnosis.
Step 7: Clinical Review
A qualified healthcare professional reviews the prediction together with the patient's clinical information before making a medical decision.

C. System Workflow



D. Advantages of the Proposed System
The proposed system offers several potential advantages:
It can assist in analysing large amounts of medical information.
It can help identify patterns in medical images and structured data.
It may support the prioritisation of cases that require further review.
It can provide consistent predictions when the same validated model and inputs are used.
It may help reduce repetitive analysis tasks.
It can support clinical decision-making when used with appropriate supervision.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 6, October 2026
Impact Factor: 8.2

IV. ALGORITHMS AND METHODOLOGY

Artificial Intelligence includes several algorithms that can be used for different medical diagnostic tasks. The choice of algorithm depends on the type of data, dataset size, clinical objective, interpretability requirements, and consequences of incorrect predictions.

Logistic Regression

Logistic Regression is a statistical learning technique used to estimate the probability of an outcome belonging to a particular category. In healthcare applications, it can examine variables such as patient age, symptoms, and laboratory measurements to estimate the possibility of a medical condition.

The resulting probability can help healthcare professionals identify cases that may require additional examination. However, this estimate should be treated as supporting information rather than a final medical conclusion. Due to its relatively simple structure, Logistic Regression is also useful as a reference model when comparing different prediction techniques.

Decision Tree

A Decision Tree organises information into a branching structure in which individual decisions depend on selected characteristics of the input data. Each branch represents a possible decision path, while the final node indicates the predicted category or outcome.

This method is useful in medical applications because its decision paths can be examined more easily than those of many complex models. However, a tree with excessive depth may learn specific details of the training dataset instead of general patterns. Techniques such as pruning and validation can help reduce this problem.

Random Forest

Random Forest is an ensemble learning approach that combines the results of several decision trees to produce a final prediction. During model development, the trees learn from different samples or subsets of available features. Combining multiple trees can improve prediction stability and help the model recognise complicated relationships in medical datasets. This technique may be useful for analysing structured patient information. Nevertheless, understanding the contribution of individual trees can be more difficult than interpreting a single decision tree.

Support Vector Machine

Support Vector Machine (SVM) is a supervised learning method that separates data into categories by determining a suitable decision boundary. Its performance depends on the characteristics of the dataset and the parameters selected during training.

In medical diagnosis research, SVM can be applied to patient records, extracted image features, and other numerical representations of clinical information. Kernel functions allow the method to handle certain nonlinear relationships. Appropriate parameter selection and independent testing are necessary to determine whether the model is suitable for a particular diagnostic task.

Artificial Neural Networks

Artificial Neural Networks (ANNs) are computational models made up of interconnected processing units arranged in layers. These layers transform the supplied information and learn relationships that may be useful for prediction. In healthcare, neural networks can be applied to medical images, physiological signals, and clinical data. Deep Learning extends this approach by using networks with multiple processing layers to learn more detailed representations of input information.

Although these methods can identify complex patterns, their development may require substantial training data and computational resources. Their predictions should therefore be evaluated carefully, and suitable methods should be used to understand their limitations.

Copyright to IJARSCT
www.ijarsct.co.in
DOI: 10.48175/568

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 6, October 2026
Impact Factor: 8.2

Comparison of Algorithms

Algorithm	Main application	Key advantage	Limitation
Logistic Regression	Risk estimation	Comparatively interpretable	Limited without nonlinear extensions
Decision Tree	Classification	Easy to visualise	May overfit
Random Forest	Classification	Combines multiple trees	Less transparent
SVM	Classification	Effective in some high-dimensional settings	Parameter selection can be difficult
Neural Network	Complex pattern recognition	Learns complex relationships	Requires careful training and validation

Data pre-processing

Medical datasets may contain incomplete entries, repeated observations, inconsistent units, or errors introduced during data collection. Such problems can influence the reliability of an AI model. Therefore, the collected information should be reviewed and prepared before it is used for model development.

The preparation process may include the following activities:

Identifying and handling repeated records.

Examining incomplete entries and selecting suitable methods to address missing information.

Converting measurements into consistent units.

Correcting formatting problems and checking unusual values.

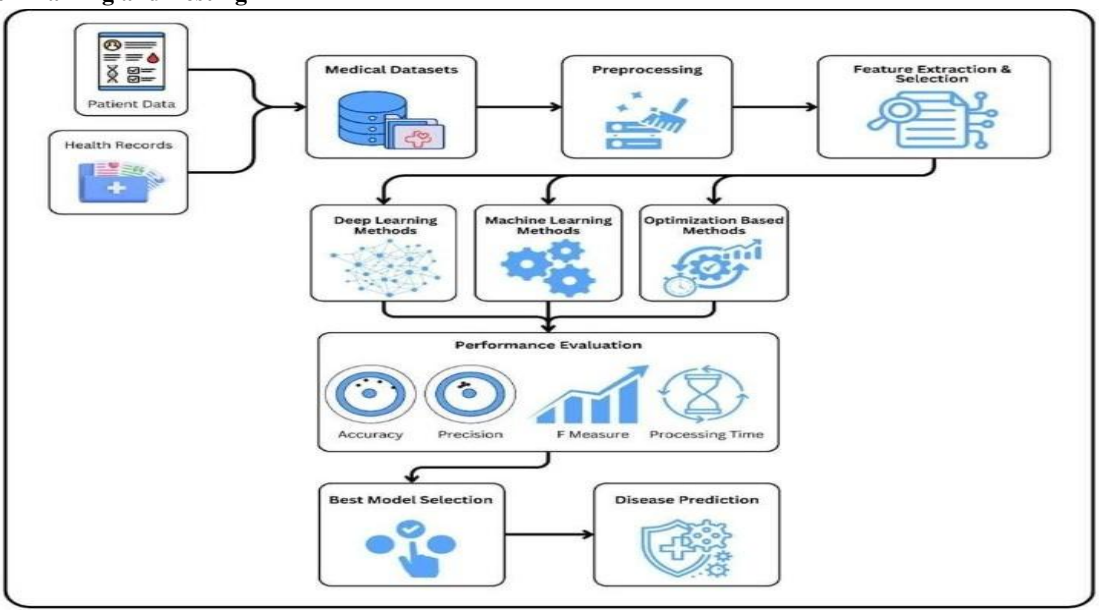
Converting categorical information into a form that the selected algorithm can process.

Preparing medical images or clinical text according to the model's requirements.

Dividing the dataset into training, validation, and testing groups.

The testing group should remain separate from the data used to train or tune the model. This separation helps provide a more independent assessment of how well the system may perform on previously unseen information.

Model Training and Testing



IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 6, October 2026
Impact Factor: 8.2

V. APPLICATIONS OF AI IN MEDICAL DIAGNOSIS

Artificial Intelligence can assist with several healthcare activities by examining medical information and identifying patterns that may be relevant to clinical assessment. The usefulness of each application depends on the quality of its input data, the reliability of the model, and the environment in which it is used.

Medical Image Analysis

AI techniques can examine radiological images, including X-rays, CT scans, and MRI scans, to identify visual features associated with particular medical conditions. Image-analysis models may help highlight areas that deserve closer examination by a radiologist.

These tools can support the review and prioritisation of medical images, especially when healthcare professionals must examine many cases. However, performance must be assessed across different imaging devices, hospitals, and patient groups before the system is used in clinical practice.

Disease-Risk Prediction

Machine Learning methods can combine information such as symptoms, previous medical conditions, laboratory results, and patient characteristics to estimate the likelihood of a particular disease.

Such estimates may help healthcare professionals decide whether additional investigations or closer monitoring are appropriate. The reliability of the output depends on the training data and the suitability of the model for the patient population. A calculated risk should not be treated as proof that a disease is present or absent.

Laboratory Report Analysis

AI systems can assist in examining laboratory measurements and identifying combinations of values that may require further attention. They can help organise complex results and highlight patterns that might otherwise need additional manual review.

Laboratory findings must still be interpreted alongside symptoms, medical history, current health status, and appropriate reference ranges. The system should also account for missing information and differences in laboratory procedures.

Electronic Health Records

Electronic Health Records (EHRs) contain information such as diagnoses, prescribed medicines, laboratory findings, and clinical notes. AI techniques can help organise these records and identify relationships that may be relevant to patient assessment.

For these applications to be useful, the information must be accurate, appropriately formatted, and handled under suitable privacy and security controls. Access to patient information should be limited to authorised purposes.

Remote Healthcare

AI-assisted tools may help healthcare services review incoming patient information and identify cases that could require timely attention. Such tools may be particularly useful where access to specialist services is limited.

Before implementation, the system should be evaluated for the intended population, language, and healthcare setting. Its output must not create unnecessary reassurance or delay appropriate medical assessment.

VI. CHALLENGES AND LIMITATIONS

Despite its potential benefits, Artificial Intelligence faces important challenges in medical diagnosis.

A. Data Privacy

Medical information is sensitive. Healthcare systems must protect patient data through appropriate access controls, secure storage, encryption, and compliance with applicable privacy requirements.

B. Data Quality

Missing values, incorrect labels, inconsistent measurements, and duplicated records can reduce model reliability. A model trained on poor-quality data may produce misleading predictions. Therefore, data collection and validation are essential parts of development.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 6, October 2026
Impact Factor: 8.2

C. Bias and Fairness

If certain patient groups are under-represented in the training dataset, the model may perform differently for those groups.

Developers should examine performance across relevant patient populations and investigate differences in errors.

D. Explainability

Some complex AI models do not provide straightforward explanations for their predictions. Healthcare professionals need enough information to judge whether an output is relevant and consistent with other clinical evidence.

E. Clinical Validation

A model that performs well on one dataset may not perform equally well in another hospital or population. Clinical validation, appropriate testing, and ongoing monitoring are necessary before a system is used in routine healthcare.

F. Human Oversight

AI predictions can be incorrect. Healthcare professionals must review the output alongside patient information and other evidence.

The system should support clinical judgement rather than replace it.

VII. DISCUSSION

The proposed Medical Diagnosis System demonstrates how AI methods can be organised into a workflow for clinical decision support.

Different algorithms offer different advantages. Logistic Regression provides a comparatively interpretable baseline, Decision Trees offer a visual decision structure, and Random Forest combines multiple trees. Support Vector Machines can be useful for certain classification tasks, while Neural Networks can learn complex representations from images, signals, and text.

However, algorithm selection should depend on the clinical problem rather than the complexity of the model. A more complex model is not automatically safer or more useful.

The quality of training data, the definition of the target condition, the evaluation method, and the clinical environment all influence system performance.

A successful medical AI system must also consider usability, data privacy, explainability, and the consequences of incorrect predictions. Evaluation should measure not only statistical performance but also whether the system supports an appropriate clinical workflow.

For these reasons, AI should be introduced through a controlled process involving technical development, clinical expertise, validation, and monitoring.

VIII. FUTURE SCOPE

AI-based medical diagnosis has several opportunities for further development.

Multimodal AI: Combining medical images, laboratory data, and clinical notes.

Explainable AI: Developing methods that make model outputs easier to review.

Privacy-Preserving Learning: Exploring approaches that protect patient information during model development.

Improved Data Quality: Developing methods to handle missing and inconsistent medical records.

External Validation: Testing models across different hospitals and patient populations.

Continuous Monitoring: Identifying changes in data and model performance after deployment.

Clinical Workflow Integration: Designing systems that support healthcare professionals without creating unnecessary alerts.

Responsible Generative AI: Exploring clinical-document summarisation with appropriate verification and safeguards.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 6, October 2026
Impact Factor: 8.2

Future research should focus on demonstrating reliable performance and meaningful clinical benefits using suitable datasets and evaluation procedures.

IX. CONCLUSION

This research paper presented the principles, architecture, algorithms, applications, and challenges of a Medical Diagnosis System Using Artificial Intelligence.

The proposed framework includes data collection, pre-processing, feature extraction, model training, prediction, performance evaluation, and clinical review. Machine Learning and Deep Learning methods can assist in analysing medical information and identifying patterns that may be relevant to diagnosis.

The paper also discussed evaluation metrics such as accuracy, precision, sensitivity, specificity, and F1-score. These measures help assess model performance, but they must be interpreted according to the clinical task and consequences of diagnostic errors.

Important challenges include data privacy, data quality, bias, explainability, generalisation, and clinical validation. Addressing these issues is necessary for the safe and responsible use of AI in healthcare.

In conclusion, Artificial Intelligence has the potential to support healthcare professionals by improving the analysis of medical information. However, AI-generated predictions should be reviewed by qualified professionals, and real-world use should follow appropriate validation, safety, and governance procedures.

ACKNOWLEDGMENT

The author acknowledges the academic resources and educational guidance that support the study of Artificial Intelligence and healthcare technology. This research paper has been prepared as an educational, literature-based study of AI-assisted medical diagnosis.

REFERENCES

[1] World Health Organization, Ethics and Governance of Artificial Intelligence for Health: WHO Guidance. Geneva, Switzerland: WHO, 2021.

[2] World Health Organization, Regulatory Considerations on Artificial Intelligence for Health. Geneva, Switzerland: WHO, 2023.

[3] A. L. Beam and I. S. Kohane, "Big data and machine learning in health care," JAMA, vol. 319, no. 13, pp. 1317–1318, 2018.

[4] E. J. Topol, "High-performance medicine: the convergence of human and artificial intelligence," Nature Medicine, vol. 25, pp. 44–56, 2019.

[5] D. S. Kermany et al., "Identifying medical diagnoses and treatable diseases by image-based deep learning," Cell, vol. 172, no. 5, pp. 1122–1131.e9, 2018.

[6] A. Esteva et al., "A guide to deep learning in healthcare," Nature Medicine, vol. 25, pp. 24–29, 2019.

[7] A. Rajkomar, J. Dean, and I. Kohane, "Machine learning in medicine," New England Journal of Medicine, vol. 380, pp. 1347–1358, 2019.

[8] X. Liu et al., "Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension," Nature Medicine, vol. 26, pp. 1364–1374, 2020.

[9] R. Miotto, F. Wang, S. Wang, X. Jiang, and J. T. Dudley, "Deep learning for healthcare: review, opportunities and challenges," Briefings in Bioinformatics, vol. 19, no. 6, pp. 1236–1246, 2018.

[10] U.S. Food and Drug Administration, Health Canada, and Medicines and Healthcare products Regulatory Agency, Good Machine Learning Practice for Medical Device Development: Guiding Principles, 2021.