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 5, May 2026
Impact Factor: 8.2

MediForecast: Multiple Disease Prediction Using Machine Learning

Badhekar Shekhar Ankush[1], Kapadi Sandesh Hanuman[2], Bombe Shubham Shankar[3], Londhe Sameer Ananda[4]
[1234]Department of Computer Science,
Samarth College of Computer Science, Belhe
shekharankush01@gmail.com[1], sandeshkapadi02@gmail.com[2],
shubhambombe03@gmail.com[3], sameerlondhe04@gmail.com[4]

Abstract: *MediForecast: Multiple Disease Prediction is a data-driven healthcare analytics system designed to predict the probability of multiple chronic diseases within a unified framework. Chronic diseases such as Diabetes, Cardiovascular Disease, Parkinson's Disease, and Kidney Disorders are increasing globally and require early detection for effective treatment. Traditional diagnostic methods are time-consuming and dependent on clinical expertise, often delaying early-stage identification.*
Multiple machine learning algorithms including Logistic Regression, Random Forest, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Artificial Neural Networks (ANN) are implemented and evaluated. The system uses a comparative approach to identify the most accurate model for each disease. Performance is evaluated using metrics such as accuracy, precision, recall, F1-score, and ROC-AUC.
This research proposes a machine learning-based system that analyzes patient health data including demographic details, medical history, and lifestyle factors. Various algorithms such as Logistic Regression, Random Forest, Support Vector Machine, and Artificial Neural Networks are used to predict multiple diseases simultaneously. The system applies preprocessing techniques like normalization, feature selection, and class balancing to improve prediction accuracy.

Keywords: Machine Learning, Disease Prediction, Healthcare Analytics, Multi-Disease System, Artificial Intelligence

I. INTRODUCTION

Healthcare is rapidly evolving with the integration of Artificial Intelligence (AI) and Machine Learning (ML). With the growth of digital medical data, these technologies have become essential tools for improving diagnosis and treatment. Chronic diseases such as diabetes, heart disease, and kidney disorders are increasing due to lifestyle changes and aging populations.

Traditional healthcare systems rely heavily on manual diagnosis, clinical testing, and expert interpretation. While these methods are reliable, they are not always efficient, especially in regions with limited medical resources. Moreover, diagnosing multiple diseases simultaneously is complex and requires significant time and effort.

Traditional diagnostic systems rely on laboratory tests and expert analysis, which can be time-consuming and expensive. Additionally, early symptoms of many diseases are subtle and difficult to detect manually. This creates a need for an intelligent system that can analyze large amounts of health data and provide early predictions.

The proposed system, MediForecast, aims to address these challenges by providing a unified platform for predicting multiple diseases simultaneously using machine learning techniques.

The MediForecast system is designed to provide a unified platform that predicts multiple diseases simultaneously using machine learning techniques. This reduces redundancy and improves efficiency compared to single-disease prediction 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 5, May 2026
Impact Factor: 8.2

II. OBJECTIVES

- The main objective is to develop a multi-disease prediction system using machine learning that can predict multiple diseases on a single platform, improving efficiency and reducing time.
- The study aims to implement and compare different machine learning algorithms to identify the most accurate and reliable model for disease prediction
- The research focuses on applying data preprocessing techniques such as cleaning, normalization, and handling missing values to improve data quality.
- The system uses feature selection methods to identify important health parameters, helping to improve model accuracy and reduce complexity.
- Another objective is to enhance prediction performance using optimization and model tuning techniques for better results.
- The system aims to provide real-time predictions and support early disease detection for better preventive healthcare decision-making

III. SCOPE

The scope of this research is focused on the application of machine learning techniques for predicting multiple diseases within a single integrated system. The study aims to provide a unified framework that improves efficiency and reduces redundancy compared to traditional diagnostic approaches

The system considers various healthcare parameters such as age, gender, body mass index (BMI), blood pressure, glucose level, cholesterol level, and lifestyle habits. These parameters are used to analyze patterns in patient data and generate accurate predictions

The scope also includes essential processes such as data preprocessing, data cleaning, normalization, and feature selection. These steps help in improving data quality and enhancing the overall performance of machine learning models

The research involves the development, training, and evaluation of multiple machine learning models. Performance is analyzed using standard evaluation metrics to ensure reliability and effectiveness of the prediction system

The system is designed as a prototype-level web-based application that demonstrates real-time disease prediction capabilities. It provides a practical understanding of how machine learning can be applied in healthcare systems

However, the scope of the study is limited to analytical and predictive modeling. It does not include clinical validation, real-time hospital integration, or deployment in large-scale medical environments

IV. LITERATURE SURVEY

SR. NO	TITLE	YEAR	AUTHER	SUMMARY
1.	Machine Learning-Based Early Detection of Diabetes	2019	Sharma et al.	This study focuses on the use of machine learning algorithms for early detection of diabetes. Decision Tree and Logistic Regression models were used to analyze patient data, Achieving high Accuracy and demonstrating the effectiveness of ML in identifying early- stage diabetes
2.	Intelligent Heart Disease Prediction Using Hybrid Machine Learning Models	2020	Kumar & Singh	This research proposes a hybrid machine learning approach combining multiple algorithms such as Random Forest and SVM to improve the accuracy of heart disease prediction. The study high lightsthe benefits of ensemble techniques in healthcare analytics.

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 5, May 2026 Impact Factor: 8.2

3.	Parkinson's Disease Detection Using Voice and Movement Analysis	2021	Ahmad et al.	The paper presents a system for detecting Parkinson's disease using voice signals and movement data. Machine learning Techniques are applied to identify patterns in speech and motor functions, enabling early diagnosis.
4.	Multi-Disease Prediction Using Integrated Machine Learning Models	2022	Rao & Verma	This study introduces a framework for predicting multiple diseases simultaneously using integrated ML models. It reduces redundancy and improves efficiency compared to single-disease prediction systems.

V. METHODLOGY

This research study focuses on developing a machine learning-based system for predicting multiple diseases using healthcare data. The study follows a structured and analytical approach to understand how different machine learning techniques can be applied to improve disease prediction accuracy and support healthcare decision-making. The research is based on the collection of data from secondary sources such as publicly available healthcare datasets, research papers, and online repositories

A comprehensive data collection process was carried out using datasets obtained from sources such as Kaggle and the UCI Machine Learning Repository. These datasets include important health-related parameters such as age, gender, body mass index, blood pressure, glucose level, cholesterol level, and lifestyle habits. The collected data was analyzed to ensure its relevance and suitability for disease prediction

The next step in the methodology involves data preprocessing, which plays a crucial role in improving the quality of the dataset. Various preprocessing techniques such as handling missing values, removing noise and outliers, normalization, and encoding of categorical variables were applied. These steps help in preparing the data for effective model training and improve the overall performance of the system

Feature selection techniques were also applied to identify the most relevant attributes that contribute significantly to disease prediction. Methods such as Recursive Feature Elimination and feature importance analysis using Random Forest were used to reduce dimensionality and enhance model efficiency

VI. SYSTEM ARCHITECTURE

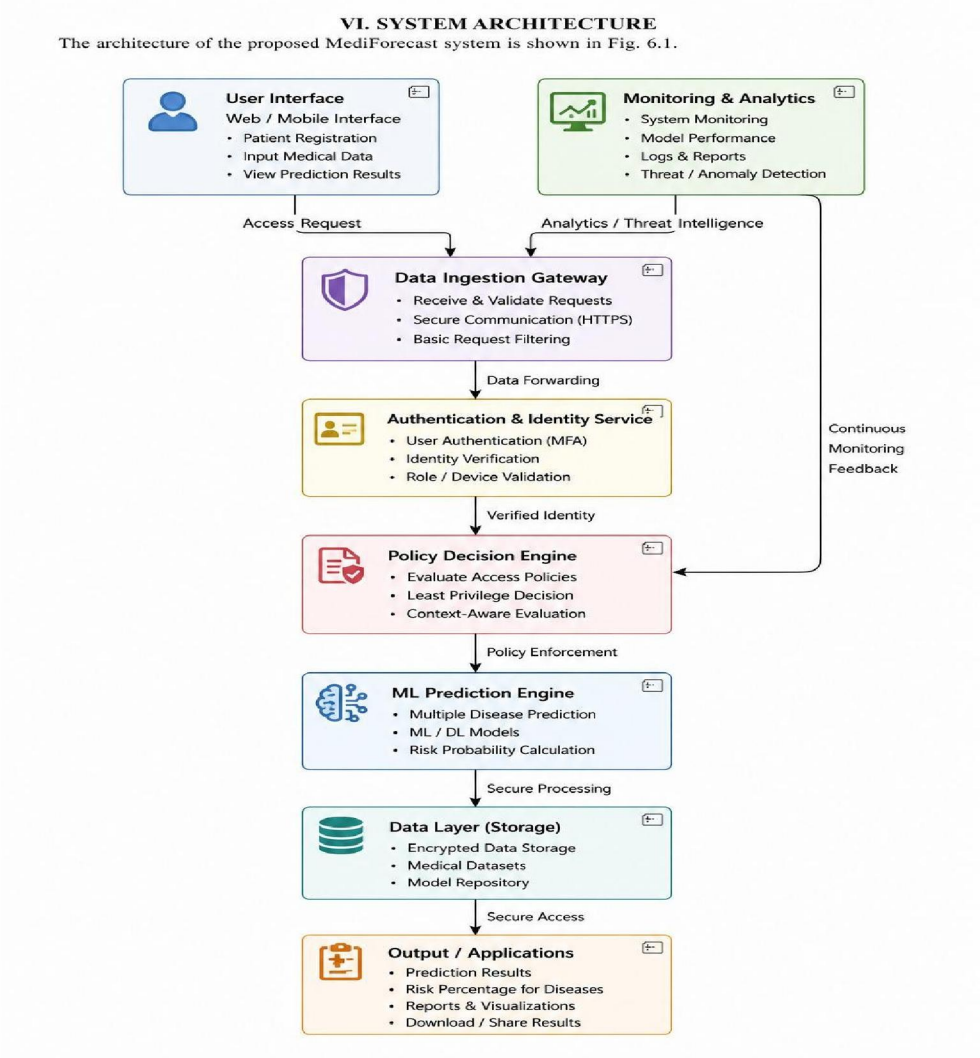
Overview of System Architecture

The system architecture of the Medi Forecast: Multiple Disease Prediction System is designed to provide an efficient and intelligent framework for predicting multiple diseases using machine learning techniques. Unlike traditional healthcare systems that rely on manual diagnosis, the proposed system follows a data-driven approach where patient information is continuously processed and analyzed to generate accurate predictions

In this architecture, the process begins with the User Interface, where users enter their health-related data such as age, blood pressure, glucose level, cholesterol, and lifestyle habits. The input data is then forwarded to the Data Preprocessing Module, which cleans and transforms the raw data into a suitable format for analysis

After preprocessing, the data is passed to the Feature Selection Module, where the most relevant attributes are selected for prediction. The refined data is then processed by multiple Machine Learning Models, which analyze patterns and generate predictions for different diseases. Finally, the prediction results are displayed through the Output Interface, providing users with disease risk levels and relevant insights

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 5, May 2026
Impact Factor: 8.2



6.2 Architectural Components and Hardware Components

User devices such as laptops, desktops, smartphones, and tablets are used to input health-related data and access prediction result

Server systems are used to store datasets and run machine learning models efficiently for processing large amounts of data

Network infrastructure devices such as routers and switches are used to ensure smooth communication between users and the system

Storage devices and databases are used to securely store medical data, trained models, and prediction results

6.3 Architectural Models

- **Data Input Model**
  The process begins with the user entering medical and health-related data through the system interface. This data acts as the input for the prediction system

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 5, May 2026
Impact Factor: 8.2

- **Data Processing Model**
  The input data is processed using techniques such as data cleaning, normalization, and transformation to make it suitable for analysis
- **Prediction** Model
  Machine learning algorithms analyze the processed data and identify patterns to generate predictions for different diseases
- **Multi-Disease Output Model**
  The system provides prediction results for multiple diseases along with risk levels, helping users understand their health condition

6.4 Case Studies or Examples

Case Study 1: AI-Based Disease Prediction Systems

In modern healthcare, machine learning-based systems are widely used to predict diseases such as diabetes and heart conditions. These systems analyze patient data and provide early warnings, helping in preventive healthcare and improving treatment outcomes

Case Study 2:Multi-Disease Prediction Systems

Recent research has introduced systems that can predict multiple diseases simultaneously using integrated machine learning models. These systems reduce redundancy, improve efficiency, and provide comprehensive health analysis compared to traditional single-disease prediction methods

6.5 Future Trends

- Integration of Artificial Intelligence and Deep Learning
  Future systems will use advanced AI and deep learning techniques to improve prediction accuracy and identify complex patterns in healthcare data
- Real-Time Health Monitoring
  Integration with wearable devices will enable continuous health monitoring and real-time disease prediction for better preventive care
- Cloud-Based Healthcare Systems
  Cloud technology will allow scalable and efficient storage and processing of large healthcare datasets, improving system accessibility
- Automated Decision Support Systems
  Future systems will provide automated recommendations and assist healthcare professionals in making faster and more accurate decisions

VII. FINDINGS

The findings of this research highlight the effectiveness of the MediForecast system in improving disease prediction and healthcare decision-making using machine learning techniques. The study demonstrates how different models, data processing methods, and system design contribute to accurate and reliable prediction results

Improved Prediction Accuracy

The study shows that machine learning algorithms significantly improve the accuracy of disease prediction compared to traditional diagnostic methods. Models such as Random Forest and Support Vector Machine provide better performance in identifying disease patterns

Multi-Disease Prediction Capability

The system is capable of predicting multiple diseases simultaneously using a single dataset. This reduces redundancy and provides a comprehensive health analysis

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 5, May 2026
Impact Factor: 8.2

Importance of Data Preprocessing

Data preprocessing plays a crucial role in improving model performance. Techniques such as normalization, handling missing values, and noise removal enhance the quality of input data

Feature Selection Enhances Performance

Selecting relevant features helps in improving model accuracy and reducing computational complexity. It also prevents overfitting and improves reliability

VIII. DISCUSSION

The first major discussion point focuses on the effectiveness of machine learning techniques in improving healthcare prediction systems. The study clearly demonstrates that machine learning models are capable of analyzing large and complex healthcare datasets to identify hidden patterns that are difficult to detect using traditional diagnostic methods. This leads to improved accuracy and faster prediction results. The use of multiple algorithms further enhances system performance by allowing comparison and selection of the most suitable model. Additionally, the integration of preprocessing and feature selection techniques ensures that the models are trained on high-quality data, resulting in more reliable and consistent predictions

The second important discussion point highlights the advantages and challenges of implementing a multi-disease prediction system. The proposed system provides the ability to predict multiple diseases simultaneously, which improves efficiency and reduces redundancy compared to single-disease systems. This makes it more practical for real-world healthcare applications. However, the study also identifies certain limitations such as the lack of clinical validation, dependency on dataset quality, and challenges in integrating the system with existing hospital infrastructure. Issues related to data privacy, security, and real-time deployment must be addressed in future research to ensure successful implementation in real-world environments

IX. CONCLUSION

MediForecast: Multiple Disease Prediction System has emerged as an effective and intelligent solution for addressing the challenges of early disease detection in modern healthcare systems. Unlike traditional diagnostic approaches that rely heavily on manual analysis and isolated testing procedures, the proposed system utilizes machine learning techniques to analyze patient data and generate accurate predictions for multiple diseases simultaneously.

This research highlights that key components such as data preprocessing, feature selection, model optimization, and multi-algorithm integration play a significant role in improving prediction accuracy and system efficiency. The study demonstrates that machine learning models can successfully identify hidden patterns in medical data, enabling early detection of diseases such as diabetes, heart disease, kidney disorders, and hypertension.

The findings also show that the system provides a scalable and user-friendly platform that supports real-time prediction and decision-making. By integrating multiple models into a unified framework, the system reduces redundancy and enhances overall performance compared to traditional single-disease prediction systems. Furthermore, the use of visualization techniques improves the interpretability of results, making the system more accessible to both healthcare professionals and general users.

In addition, the proposed system is well suited for modern healthcare environments, including digital health platforms, remote monitoring systems, and cloud-based applications. Although the implementation of such systems may require high-quality datasets, proper model validation, and integration with existing healthcare infrastructure, the long-term benefits include improved diagnostic efficiency, reduced healthcare costs, and better patient outcomes

REFERENCES

[1]. J. Smith and A. Kumar, “Machine Learning Approaches for Disease Prediction,” *International Journal of Computer Applications*, vol. 178, no. 40, pp. 1–6, 2019.

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 5, May 2026 Impact Factor: 8.2

[2]. P. Sharma, R. Gupta, and S. Verma, "Early Detection of Diabetes Using Machine Learning Techniques," *Procedia Computer Science*, vol. 167, pp. 161–170, 2020.
[3]. A. Kumar and R. Singh, "Intelligent Heart Disease Prediction Using Hybrid Machine Learning Models," *International Journal of Engineering Research & Technology*, vol. 9, no. 5, pp. 234–239, 2020.
[4]. M. Ahmad, S. Khan, and T. Ali, "Parkinson's Disease Detection Using Voice and Machine Learning Techniques," *IEEE Access*, vol. 9, pp. 112345–112356, 2021.
[5]. S. Rao and P. Verma, "Multi-Disease Prediction System Using Machine Learning Algorithms," *International Journal of Advanced Research in Computer Science*, vol. 13, no. 2, pp. 45–52, 2022.
[6]. R. Deshpande, M. Kulkarni, and S. Patil, "Machine Learning Framework for Preventive Healthcare Analytics," *Journal of Healthcare Engineering*, vol. 2020, Article ID 123456, 2020.
[7]. T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD Conference*, 2016, pp. 785–794.
[8]. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.
[9]. F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
[10]. J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed., Morgan Kaufmann, 2011.
[11]. World Health Organization, "Global Health Estimates on Chronic Diseases," WHO Report, 2021. [12]. Kaggle, "Medical Dataset for Disease Prediction," Available: https://www.kaggle.com
[13]. UCI Machine Learning Repository, "Healthcare Datasets," University of California, Irvine, 2022.
[14]. S. Haykin, *Neural Networks and Learning Machines*, 3rd ed., Pearson, 2009.