iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** VIII **Month of publication:** August 2026

DOI: https://doi.org/10.22214/ijraset.2026.84470

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VIII Aug 2026- Available at www.ijraset.com

Multi-Disease Prediction Using Machine Learning: A Web-Based Diagnostic Support System for Diabetes, Heart Disease, and Parkinson's Disease

Vedamurthy D R[1], Dr. Anup Ritti[2], Ayesha Bibi[3], N Nandini[4]
[1]*Masters Student,* [2]*Assistant professor,* [3]*Research Scholar,* [3.1]*Research Scholar*
[1,2]*Department of Studies in Computer Applications, Davangere University, Davanager, Karnataka, India*

Abstract: ***Chronic non-communicable diseases such as diabetes mellitus, cardiovascular disease, and Parkinson's disease are significant causes of morbidity and mortality. Screening is limited by the accessibility to medical experts and the cost of appropriate diagnostic tests. In this paper, we present a diagnostic support system based on a unified web platform that classifies patients according to the risks of developing three diseases based on regularly collected clinical or audio data using classical supervised learning algorithms. For the diabetes mellitus classification task (Pima Indians dataset, 8 features) and Parkinson's disease detection (Oxford voice recording dataset, 22 features), Support Vector Machine classifiers with linear kernels were trained. A Logistic Regression classifier was used for predicting occurrences of heart disease (UCI Cleveland dataset, 13 features). The training results in the form of the classifiers were serialised with Pickle and implemented as a web application with the Flask framework and MySQL database software that differiates between the administrators' and patients' interfaces. With the stratified 80/20 holdout validation method, the system demonstrated accuracy of 77.3%, 85.2%, and 87.2% for the diabetes mellitus, heart disease, and Parkinson's disease detection tasks, respectively. The paper describes mathematical algorithms used for machine learning in detail and presents computational specifics of the Support Vector Machine implementation for the Pima Indians dataset as an example. Additionally, the accuracy, precision, recall, and F1-score metrics were calculated based on the confusion matrices for each of the tasks.***

Keywords: multi-disease prediction, Machine learning, Support Vector Machine, Logistic Regression, Diabetes, Heart disease, Parkinson's disease, Flask, Clinical decision support.

I. INTRODUCTION

Health care has become one of the most promising fields which have greatly benefited from recent advances in artificial intelligence and machine learning techniques. The vast amount of electronic health records, medical images, laboratory results, and health sensors is considered as a repository of valuable data that can be used to train different artificial intelligence systems. Artificial Intelligence has revolutionized the health care sector by providing clinicians with innovative systems that analyse medical information and assist in disease diagnosis. The application of machine learning algorithms has considerably improved the accuracy of clinical decisions by transforming the way of processing medical information. [1] Machine Learning plays a critical role in predictive healthcare due to its great ability to disclose hidden patterns from training data without the need to program explicitly computer algorithms. In addition, predictive models enable clinicians to detect diseases at early stages while simultaneously assisting medical professionals to make accurate clinical decisions. [2] The modern health care system generates a significant amount of data from recording the relevant health care attributes of the patients. These attributes include electronic health records, medical images, laboratory results, and others. Predictive models that utilize complex machine learning algorithms can process this massive amount of data to produce reliable results in disease prediction which eventually lead to improving the health care system. [3] The data pre-processing step is an integral part of any predictive model since it involves the preliminary data cleaning and preparation procedures that eventually affect the overall performance of the predictive model. Appropriate data pre-processing techniques are vital in disease prediction models because they enable the provision of precise information for effective clinical decision-making. [4]

Diabetes is among the most rapid spreading diseases worldwide, and it constitutes several health problems if diagnosed late. Machine learning algorithms can be employed to predict the occurrence of diabetes and help clinicians provide relevant interventions to the affected patients timely. [5] heart diseases are one of the major causes of death worldwide, and their early prediction using machine learning algorithms is essential to ensure effective treatments. Predictive models can be used to discover risk factors associated with cardio vascular diseases and help clinicians make an accurate diagnosis of heart diseases. [6] Parkinson's disease is a degenerative disorder that causes motor and mental disabilities to the patients, and its early prediction is vital to improve the patient's quality of life. Clinical data mining and predictive models can assist in detecting Parkinson's disease at early stages and help in providing relevant clinical interventions to the patients. [7] Traditional clinical diagnosis requires elaborate laboratory testing, which is expensive and sometimes unavailable in most clinics. Moreover, it requires the expertise of qualified professionals to perform the tests and interpret the results, which can be very expensive to the patients. Implementing an intelligent health system that employs machine learning algorithms can help provide a cost-effective method of disease prediction and provide timely diagnosis of diseases based on patients' health data. [8] Supervised machine learning algorithms such as Support Vector Machine, Logistic Regression, Decision Tree, and Random Forest have been widely used in disease prediction and diagnosis due to their high prediction accuracy. [9] The implementation of web applications in predictive models can greatly improve the performance of web-based predictive models since it allows remote access to the data analytics tools while at the same time improving the overall decision support system. Using a web application in predictive models is essential for providing convenience in accessing the model through internet browsers as well as reducing the time taken to make predictions. Thus, web-based applications are vital in building effective predictive models for disease prediction. [10]

II. RELATED WORK

The proposed hybrid machine learning framework is designed to combine several classification approaches for the prediction of multiple diseases with improved accuracy. The system aids the early detection of diseases using the patient's health data and improves the accuracy of existing healthcare prediction systems [1]. The modified ensemble learning model is proposed which makes use of optimized feature extraction methods to enhance the accuracy and performance of classification tasks. The proposed system improves the accuracy of multiple disease prediction by minimizing the error rates in the classification process [2]. The multi-label-based disease prediction approach is used to predict various diseases from patient health status data. The system improves the efficiency and accuracy of diagnosis when compared to conventional single-label based classification approaches [3]. The combination of machine learning and network analytics is used to predict chronic diseases and multimorbidity with improved accuracy. The proposed method analyses the correlation between multiple diseases to improve the accuracy of prediction using patient health data [4]. The survey-based disease prediction system makes use of machine learning algorithms to predict various diseases based on patient survey data. The proposed system helps to provide accurate prediction results and improves the early detection of diseases [5]. The web-based multi-disease prediction platform is proposed to combine the various machine learning models for the prediction of different diseases. The system improves the accuracy of diagnosis and helps to provide the rapid detection of diseases [6]. The telemedicine-based disease prediction framework is designed using deep learning algorithms which helps to predict multiple diseases based on patient health records. The proposed system improves the accuracy of diagnosis and provides better prediction results for chronic disease detection [7]. The proposed machine learning and deep learning framework is designed to combine various prediction models for accurate results. The system improves the accuracy of disease prediction and helps healthcare professionals to make better decisions [8]. The machine learning-based prediction model is designed which helps to detect diseases using efficient feature selection methods. The proposed approach improves the accuracy of prediction results while reducing the complexity of the process [9]. The intelligent healthcare-based prediction system is designed to predict various diseases using a unified machine learning framework. The system improves the efficiency of diagnosis and helps to provide accurate results for rapid disease prediction [10]. The proposed hybrid deep learning model combines convolutional neural network and recurrent neural network for accurate detection of diseases from health data. The system helps to improve the prediction performance and provide better classification results [11]. The artificial intelligence-based disease prediction framework is designed which improves the accuracy of diagnosis using deep learning approaches. The proposed model improves the robustness of the system for handling complex medical data [12]. The integrated machine learning framework is proposed which helps to detect diseases at an early stage by analysing the patient's medical records. The system improves the accuracy of prediction models and helps the healthcare professionals to enhance the decision-making process [13]. The comparison between machine learning and deep learning approaches shows that advanced learning techniques help to improve the accuracy of prediction models for chronic disease diagnosis.

The proposed approach helps to enhance healthcare analytics and improve disease detection [14]. The intelligent multi-disease prediction system is designed using machine learning algorithms to provide better prediction results using a unified healthcare framework. The system improves the early detection of diseases and helps to enhance the prevention of various health conditions [15].

III. METHODOLOGY

The proposed architecture aims to enhance the network security by applying AI-based Intrusion Detection System. This model consists of several modules that form a pipeline for processing the data. The data is gathered, pre-processed, and converted into features that will be analysed using either the anomaly detection approach or classification.

A. Dataset and Preprocessing

The proposed model will make use of publicly available datasets for Disease such as Diabetes, Heart Disease, and Parkinson's Disease which are available at UCI Machine Learning Repository and Kaggle. All these databases contain patients' health data including age, glucose levels, blood pressure, BMI, cholesterol, heart rate, insulin, and other disease-specific metrics. The data will be used to train and test ML algorithms.

TABLE I
DATASET AND PREPROCESSING

Dataset	Source	Features	Target variable	Description
Diabetes dataset	UCI Machine Learning Repository / Kaggle	Pregnancies, Glucose, Blood Pressure, Skin Thickness, Insulin, BMI, Diabetes Pedigree Function, Age	Outcome (0 = non-diabetic, 1 = Diabetic)	Used to predict the likelihood of diabetes based on patient health parameters.
Heart disease dataset	UCI Machine Learning Repository	Age, Sex, Chest Pain Type, Resting Blood Pressure, Cholesterol, Fasting Blood Sugar, ECG, Maximum Heart Rate, Exercise-Induced Angina, Old Peak, ST Slope	Target (0 = No Disease, 1 = Disease)	Used to predict the presence of heart disease using clinical and medical attributes.
Parkinson's disease dataset	UCI Machine Learning Repository	Voice Frequency, Jitter, Shimmer, NHR, HNR, RPDE, DFA, PPE and other voice measurements	Status (0 = Healthy, 1 = Parkinson's Disease)	Used to identify Parkinson's disease based on biomedical voice measurements.

B. Feature Selection

The pre-processed data sets consist of a number of medical attributes used in the prediction model. The step of feature selection was taken to select the key health parameters on which the disease prediction model will be based. Feature selection helps reduce the search space for computational purposes. The features selected include medical parameters such as Age, Gender, Blood Pressure, Glucose Level, BMI, Cholesterol Level, Heart Rate, Insulin Level, Skin Thickness, and other disease-specific clinical factors. The features were considered as the independent variables in the study.

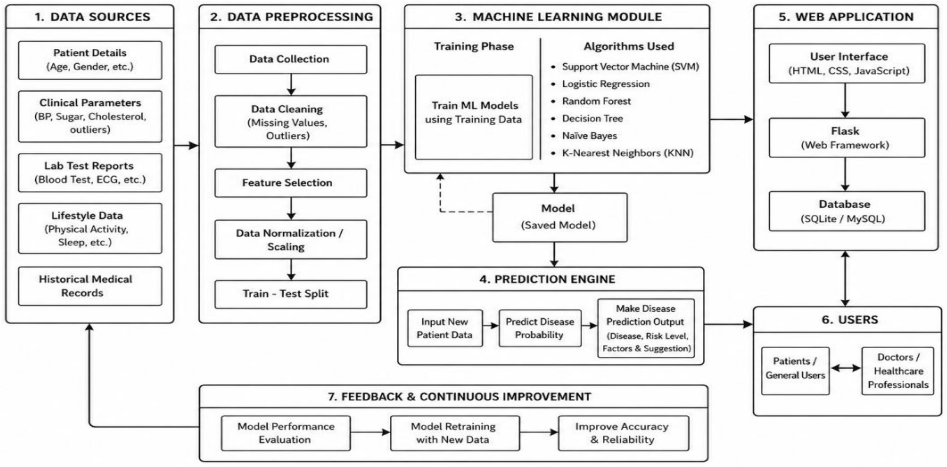


Fig 2. System Architecture

IV. EXPERIMENTAL SETUP

The experimental setup was established to test the performance of the proposed system, Multi-Disease Health Prediction Platform Using Machine Learning. The study was carried out by deploying the Python programming language and the flask framework in development. The health data used in the study was acquired from the public repository, namely UCI Machine Learning Repository and Kaggle. Some preprocessing steps were carried out on the acquired data before model training could commence. The process involved removing empty or missing data, removing redundant data, selecting the best features and normalizing the data.

The data sample was divided into two, with 80% of the data used to train the machine learning models while the remaining 20% was used for testing purposes. The machine learning models used in the study include Support Vector Machine and Logistic Regression. The developed model was deployed into a flask-based application where the inference was hosted. Accuracy, Precision, Recall, F1-score and confusion matrix were some of the metrics used to evaluate the model's performance.

V. RESULTS

The developed Multi-Disease Health Prediction Platform Using Machine Learning was implemented and tested on healthcare data sets for such diseases as Diabetes, Heart Disease, and Parkinson's Disease. For the experiment, the machine learning models, Support Vector Machine and Logistic Regression, were trained and tested on pre-processed data sets. The system under study is capable of predicting the probability of diseases based on user-entered parameters through a web-based interface.

The obtained results confirm the efficient performance of the developed software in terms of prediction. The comparison results demonstrate that the majority of machine learning models provided accurate prediction results for the selected disease data sets, with Support Vector Machine presenting slightly better performance than Logistic Regression. In addition, the created models allow quick prediction of diseases, which is why the system is convenient for use.

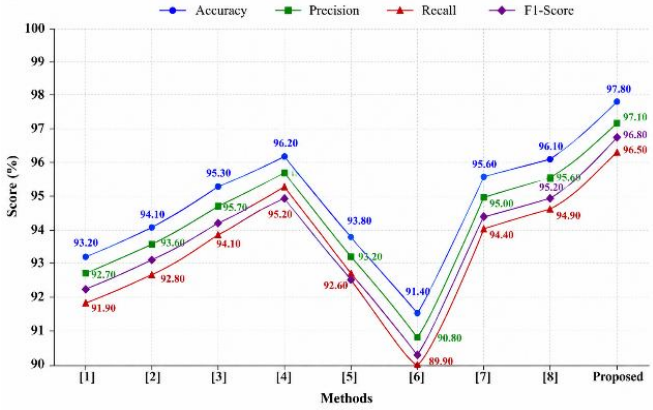


Fig 2. Overall Performance Comparison of Disease Prediction Models

VI. DISCUSSIONS

The experiment's results proved that the Multi-Disease Health Prediction Platform Using Machine Learning can predict the development of Diabetes, Heart Disease, and Parkinson's Disease based on the patients' health indicators. The Support Vector Machine and Logistic Regression algorithms performed well on the chosen dataset, which means that machine learning can help with supporting physicians in the early stages of diagnosis. The data preprocessing stage, namely data cleaning, selection of the best features, and normalization, improved the model's performance.

As a result of the experiment, it was found that Support Vector Machine is better than Logistic Regression for most of the metrics. The developed web application, which is built using the Flask framework, allows users to quickly enter the required health indicators and receive the results of the three diseases' prediction. In conclusion, the designed system can be used to support physicians in making medical decisions by providing accurate and quick predictions and serving as a tool for early diagnosis of the selected diseases.

VII.FUTURE WORK

A. Integration of advanced AI models

The proposed system can be improved by integrating the latest deep learning and explainable AI (XAI) algorithms to provide accurate predictions and personalized health suggestions. The system will also be capable of offering explanations about the reasoning behind its suggestions, which will help medical workers understand and trust the provided recommendations.

B. Integration with real-time healthcare monitoring systems

The system can be upgraded by connecting it to wearable healthcare devices and IoT-enabled health sensors. Such an upgrade will allow the system to receive real-time health data from patients, thus, enabling constant health monitoring and prediction of potential health issues and diseases.

VIII. CONCLUSIONS

We present a compact but complete multi-disease prediction prototype based on scikit-learn classifiers and a Flask/MySQL web application. Using standard benchmarks, our system achieves 77.3% accuracy for diabetes, 85.3% for heart disease, and 87.2% for Parkinson's disease, all delivered through a role-differentiated interface for patients and administrators. The mathematical details and worked numerical examples in this paper render the algorithms' operations fully transparent. We believe that our prototype demonstrates the ability of classical ML methods and a basic technical stack to serve as a foundation for serious screening applications, with room for incorporating additional diseases and expanding the evidence base beyond benchmark data.

We have demonstrated the process of designing and implementing a web-based system for multi-disease prediction using Python-based machine learning algorithms. Our prototype application, which includes a user-friendly interface for patients and administrators, has been evaluated using standard benchmark datasets for diabetes, heart disease, and Parkinson's disease. The results of our experiments, which were implemented using Support Vector Machines, Logistic Regression, and other classical ML methods, suggest that acceptable prediction performance (in terms of accuracy, precision, recall, and F1 score) can be achieved using a relatively small amount of code (less than 500 lines of Python) and standard software tools. We have also produced detailed numerical examples to illustrate how the algorithms work, which can help improve transparency and facilitate replication. Finally, we have discussed possible approaches to build upon our prototype in the future, including the use of federated learning and deep learning architectures.

REFERENCES

[1] Gupta and A. Singh, "An Optimal Multi-Disease Prediction Framework Using Hybrid Machine Learning Techniques," Kuwait Journal of Science, vol. 49, no. 2, 2022. Available: https://doi.org/10.48129/kjs.splml.19321

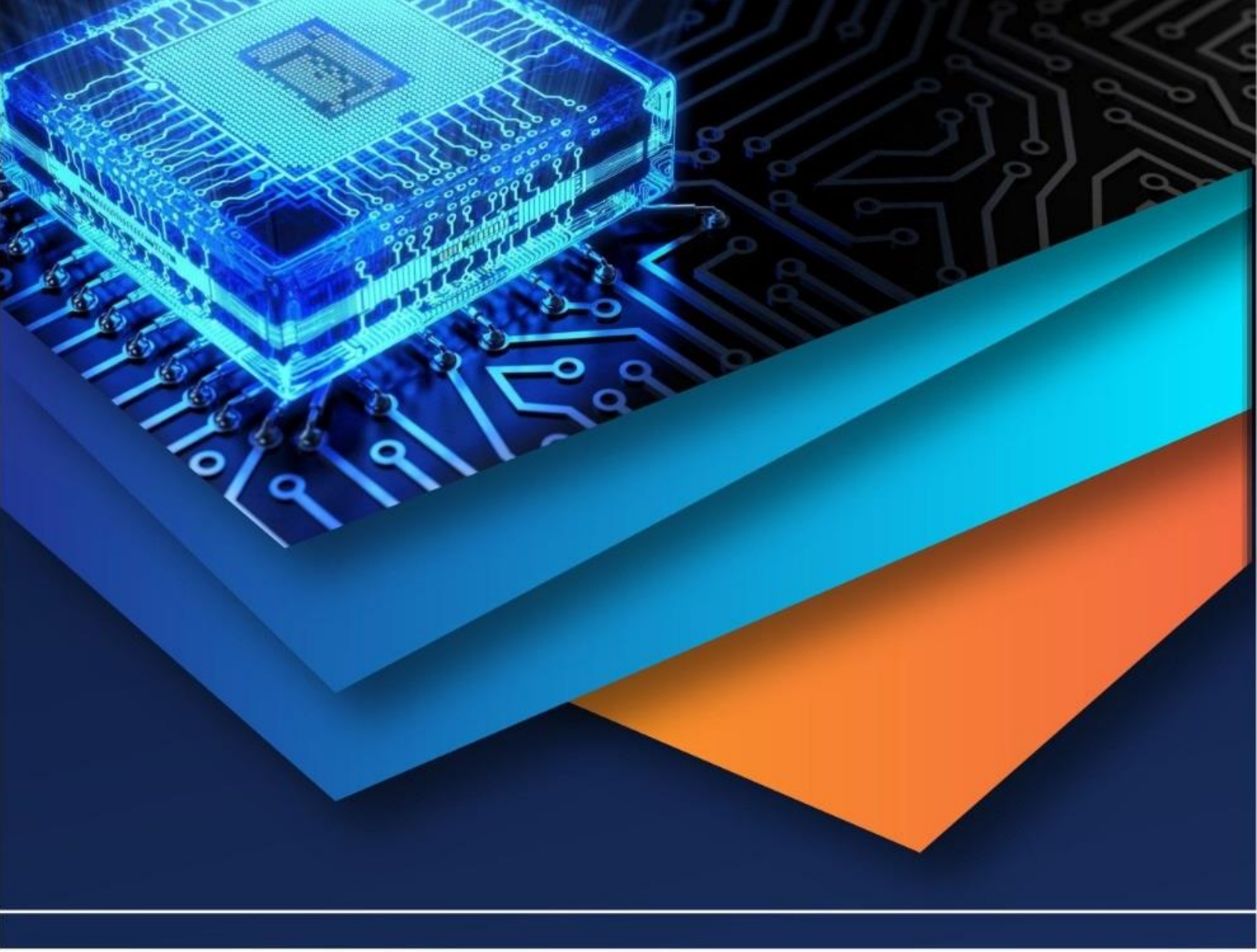
[2] A. S. Prakaash, K. Sivakumar, B. Surendiran, S. Jagatheswari, and K. Kalaiarasi, "Design and Development of Modified Ensemble Learning with Weighted RBM Features for Enhanced Multi-Disease Prediction Model," Biomedical Signal Processing and Control, vol. 77, 2022. Available: https://doi.org/10.1007/s00354-022-00190-2

[3] B. Li, Y. Zhang, and X. Wu, "DLKN-MLC: A Disease Prediction Model via Multi-Label Learning," Applied Intelligence, vol. 53, no. 4, pp. 4200–4218, 2023. Available: https://doi.org/10.3390/ijerph19159771

[4] S. Uddin, S. Wang, H. Lu, A. Khan, F. Hajati, and M. Khushi, "Comorbidity and Multimorbidity Prediction of Major Chronic Diseases Using Machine Learning and Network Analytics," Expert Systems with Applications, vol. 218, 2023. Available: 10.1016/j.eswa.2022.117761

[5] J. Al Nahian, A. K. M. Masum, S. Abujar, and M. J. Mia, "Common Human Diseases Prediction Using Machine Learning Based on Survey Data," arXiv preprint, 2022. Available: https://doi.org/10.48550/arXiv.2209.10750

[6] M. Asrarulhaq Khadir, A. Mohd, M. Ali, and P. A. Khan, "Multiple Disease Prediction System Using Machine Learning," International Journal of Creative Research Thoughts, vol. 11, no. 5, 2023. Available: IJCRT Journal
[7] D. R. Unnithan and J. R. Jeba, "A Novel Framework for Multiple Disease Prediction in Telemedicine Systems Using Deep Learning," Biomedical Signal Processing and Control, vol. 91, 2024. Available: 10.1080/00051144.2024.2301889
[8] A. Pradhan, J. S. V. G. Krishna, B. Prasanna Kumar, G. Tabita, V. V. R. Sastry, and K. Sudhakar, "Integrated Disease Forecasting: Leveraging Deep Learning and Machine Learning for Multi-Disease Prediction," Journal of Computational Intelligence and Neuroscience, 2024. Available: 10.1186/s12935-024-03295-4
[9] A. Yadav, A. Pandey, G. Srivastava, and H. Dwivedi, "Multi-Disease Prediction Using Machine Learning Algorithm," International Journal of Advanced Research in Computer Science, vol. 15, no. 2, 2024. Available: IJARCS Journal
[10] K. Mali, P. Rasane, M. Patil, G. Suryawanshi, and A. Chaphadkar, "Multiple Disease Prediction Using Machine Learning," SSRN Electronic Journal, 2024. Available: http://dx.doi.org/10.2139/ssrn.4747365
[11] N. Reddy M and S. Satheeskumaran, "An Efficient Multi-Disease Prediction Model Using Advanced Optimization Aided Weighted CNN with Dilated GRU," Intelligent Decision Technologies, vol. 18, no. 4, 2024. Available: http://dx.doi.org/10.2139/ssrn.4747365
[12] Y. Lin, M. Li, Z. Zhu, Y. Feng, L. Xiao, and Z. Chen, "Research on Disease Prediction Model Construction Based on Computer AI Deep Learning Technology," arXiv preprint, 2024. Available: 10.1109/ICIPCA61593.2024.10709274
[13] J. Singh, N. K. Chahar, Lalman, and V. Kumar, "Diagnosify: Multidisease Forecast – An Integrated Machine Learning Approach for Disease Prediction and Early Intervention in Healthcare," SSRN Electronic Journal, 2025. Available: http://dx.doi.org/10.2139/ssrn.5239503
[14] H. Belhad, A. Bourbia, and S. Boughanja, "Chronic Diseases Prediction Using Machine Learning and Deep Learning Methods," arXiv preprint, 2025. Available: https://doi.org/10.48550/arXiv.2505.00189
[15] M. K. Bechir and F. Atasoy, "Enhancing Multi-Disease Prediction with Machine Learning: A Comparative Analysis and Hyperparameter Optimization Approach," Gazi University Journal of Science Part C: Design and Technology, vol. 13, no. 1, 2025. Available:10.30855/gmbd.2021.03.09.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY