WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

SMART HEALTH ADVISOR

K. Bhuvana Satyanarayana[1],P .Arun Kumar[2], Vamsi Krishna Dandasi[3],M .Sivan Naidu[4], Shaik Sadiq Saheb[5]

[1]Dept of IT GMR Institute of Technology **kalavalapallibhuvan@gmail.com**
[2]Dept of IT GMR Institute of Technology **arunkumarperla1@gmail.com**
[3]Dept of IT GMR Institute of Technology **vamsikrishnadandasi0@gmail.com**
[4]Dept of IT GMR Institute of Technology **m.reshi333@gmail.com**
[5]dept of IT GMR Institute of Technology **sadiksaheb143@gmail.com**

ABSTRACT

This research aims to enhance the early identification of cardiovascular conditions, diabetes, and renal disorders among individuals living in remote areas with restricted access to medical services. Two machine learning techniques — Random Forest and Logistic Regression — are employed to analyze health data and uncover critical risk elements. While Random Forest is effective in detecting intricate patterns, Logistic Regression offers clear and interpretable results, balancing precision and ease of understanding. Following risk detection, suggested interventions include encouraging nutritious eating habits based on locally sourced foods, promoting consistent exercise, and facilitating easy-to-reach health screenings. Moreover, educational programs will focus on raising awareness about early intervention strategies, lifestyle modifications, and community-specific preventive actions to further benefit underserved populations.

Keywords: Cardiovascular Conditions, Diabetes, Renal Disorders, Preventive Actions, Remote Communities

INTRODUCTION

Chronic illnesses such as cardiovascular diseases, diabetes, and kidney disorders are becoming increasingly widespread and are responsible for a significant proportion of global mortality. The situation is even more critical in rural regions where limited healthcare services, delayed recognition of symptoms, and unawareness of personal risk contribute to poor management of these conditions. Early identification plays a crucial role in saving lives, but unfortunately, rural communities often lack the necessary resources and diagnostic tools.

To tackle this challenge, this study leverages machine learning methods — specifically Random Forest and Logistic Regression — to analyze healthcare data and pinpoint individuals at higher risk. These models are utilized to discover patterns and predict the likelihood of developing these chronic diseases, providing an efficient and scalable solution. The research emphasizes how predictive analytics can guide early interventions and empower healthcare initiatives to reach underserved populations more effectively.

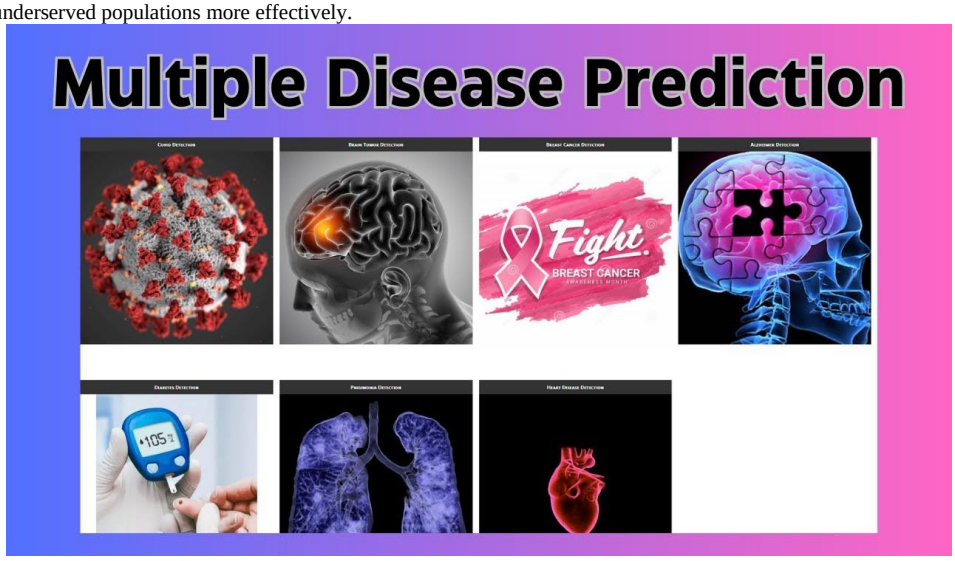


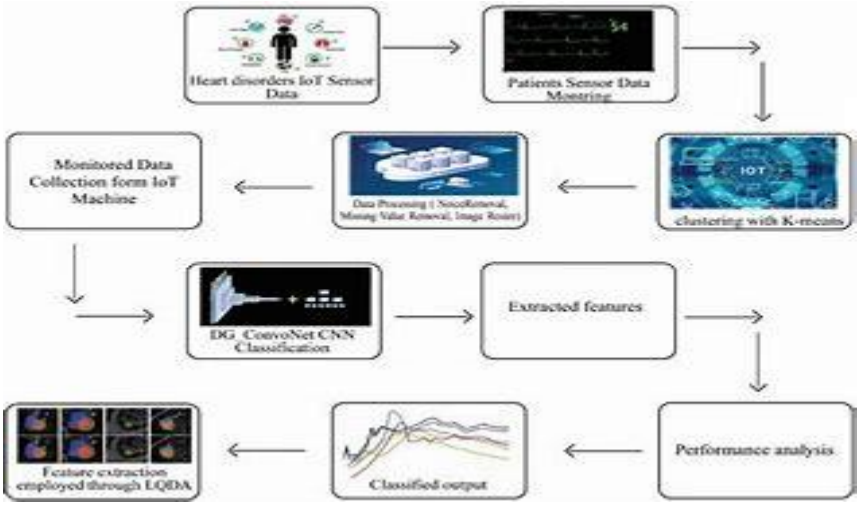
Figure1.1

Although these chronic conditions can often be detected early, which significantly improves chances of effective treatment and survival, rural residents frequently lack the necessary facilities and diagnostic technologies. This gap in early detection prevents timely medical intervention, resulting in avoidable complications and fatalities.

In response to this issue, the current study proposes the application of machine learning algorithms, specifically Random Forest and Logistic Regression models, to examine health records and recognize individuals at a higher risk. These models are capable of revealing important health patterns and offering predictive insights, thereby helping to identify threats before they escalate into severe health problems.

Alongside the technical strategies, the research advocates for the adoption of basic, actionable health practices among rural populations. Key recommendations include promoting nutrition through locally available food resources, encouraging regular physical exercise, and facilitating consistent health screenings to monitor and manage risks effectively.

Ultimately, the integration of data-driven tools with grassroots health education initiatives aims to bridge the healthcare gap in underserved regions. By strengthening early detection efforts and fostering healthier lifestyles, the study aspires to reduce the prevalence of chronic diseases and enhance the overall well-being of people living in resource-constrained areas.



LITERATURE SURVEY

[1].Machine learning has greatly advanced healthcare by allowing early identification of diseases. The Multiple Disease Prediction System (MDPS) examines patient records and lifestyle details to predict potential illnesses. It involves steps like cleaning data, selecting features, training models, and diagnosing conditions. While it supports faster diagnosis and reduces costs, issues like data security and model precision still need attention. Overall, MDPS shows strong potential for improving healthcare and managing chronic diseases efficiently.

[2]. Millions of people experience curable illnesses like heart disease, diabetes, and cancer, but poor access to affordable healthcare often results in late or wrong diagnoses. For effective treatment, diagnostic methods must be accurate, affordable, and easily available. Technologies like machine learning and deep learning are revolutionizing healthcare by improving diagnosis

speed and reliability. This work proposes building an intelligent system that uses patient records to predict multiple diseases. Such a system can minimize diagnostic errors, lower treatment expenses, and improve healthcare access in underserved regions.

[3]. Artificial Intelligence is extensively applied in healthcare to predict diseases, but conventional machine learning models often fall short in accuracy. Recurrent Neural Networks (RNNs) can track changes in medical data but have difficulties in fine-tuning their parameters, impacting their results. Techniques like Particle Swarm Optimization (PSO) enhance model performance but may not always yield optimal outcomes. Inspired by plant growth, the Artificial Flora (AF) algorithm provides a superior method for optimizing AI models. Our work presents an enhanced RNN using AF, which outperforms regular RNN and RNN-PSO models, leading to better remote healthcare solutions.

[4]. Researchers have increasingly applied machine learning (ML) and deep learning (DL) techniques to enhance early disease detection. Traditional ML models like Logistic Regression, KNN, Random Forest, and SVM are helpful for initial classification but sometimes lack precision. Deep learning approaches such as Artificial Neural Networks (ANN) can capture complex medical data patterns, boosting prediction performance. Integrating ML and DL has proven more effective for identifying diseases like cancer, diabetes, monkeypox, and Parkinson's. Our system, built with real-world datasets and Streamlit for easy use, supports doctors in faster and more accurate diagnosis.

[5]. Various studies have demonstrated the use of machine learning models like Logistic Regression and SVM in predicting diseases such as diabetes, heart disease, and Parkinson's. These models rely on health metrics like blood pressure, cholesterol, and heart rate to aid early detection. Recent AI advancements have enhanced prediction accuracy, with Streamlit enabling more interactive healthcare applications. Despite progress, most systems focus on predicting a single disease, leaving multi-disease prediction an area for further exploration.

S.no	Author	year	techniques	advantages	disadvantages	Dataset information
Paper 6	**Swaroop Sana**	**2024**	**The paper uses Logistic Regression and SVM for multi-disease prediction, utilizing health parameters with a Streamlit-based interface.**	**It offers accurate predictions, a user-friendly interface, and personalized insights, enhancing efficiency by analyzing multiple diseases on one platform.**	**Single disease prediction models can lead to fragmented analysis, missing interconnections between diseases, and ultimately resulting in delayed diagnosis and less effective treatment options for patients.**	**The dataset includes data for diabetes, heart disease, and Parkinson's, with parameters like blood pressure, pulse, cholesterol, and heart rate.**
Paper 7	**A S Prakaash**	**2022**	**The paper uses data collection, pre-processing, Deep Belief Network, ensemble learning with SVM, RNN, DNN, and ASR-CHIO optimization.**	**The model improves multi-disease prediction, aids doctor diagnoses, handles unbalanced datasets, and is suitable for real-time mobile applications.**	**Limited scalability and computational complexity are challenges that may affect the model's efficiency and performance in large-scale applications.**	**The study uses datasets for stroke prediction, heart disease, and diabetic retinopathy, with patient data ranging from 1,190 to 11,500 entries.**
Paper 8	**Aditya Gupta Amritpal Singh**	**2022**	**The paper employs hybrid machine learning techniques, including genetic algorithm-based recursive feature elimination and AdaBoost, for effective multi-disease prediction**	**The proposed methodology enhances predictive performance, reduces overfitting, and effectively handles high-dimensional data for multi-disease prediction**	**The methodology may require extensive computational resources and may not generalize well with small datasets**	**The study utilizes the Cleveland and Pima datasets, focusing on heart diseases and diabetes, which contain missing values addressed through MICE imputation**

S.no	Author	year	techniques	advantages	disadvantages	Dataset information
Paper 9	**D Dhinakaran (Corresponding Author) S Edwin Raja M Thiyagarajan J Jeno Jasmine P Raghavan**	**2024**	**The paper employs an ensemble feature selection model, the Stabilized Energy Valley Optimization with Enhanced Bounds (SEV-EB) algorithm, and an HSC-AttentionNet for disease prediction**	**The proposed model enhances predictive accuracy, achieves stability in feature selection, and captures both short-term and long-term health data patterns, significantly improving disease prediction outcomes**	**The model can be complicated, needs good quality data, raises concerns about patient privacy, and requires a lot of computer power, which might make it hard to use in real healthcare situations.**	**The study utilizes two datasets: "Covidpred," with 70,500 COVID-19 patients and 35,256 normal controls, and the "Stroke Prediction Dataset," containing 12 attributes from 11 clinical sectors to enhance disease prediction accuracy**

Paper 10	**Wei Zhang Weihan Cheng Koichi Fujiwara Richard Evans Chengyan Zhu**	**2024**	**The paper uses regression analysis, logistic regression, feature selection, EHR data extraction, and validation techniques for predicting heart disease hospital readmissions.**	**Predictive modeling enables hospitals to identify high-risk patients, improving care, reducing costs, enhancing efficiency, and supporting smarter decisions with technology.**	**Data Quality Issues Integration Challenges**	**The paper analyzes 56 studies using hospital data, administrative databases, and registries, with dataset sizes ranging from 351 to 1,420,564 records, covering an average timespan of 59 months**
Paper 11	**Dina Saif, Amany M. Sarhan, and Nada M. Elshennawy**	**2024**	**The paper employs deep learning techniques, including CNN, LSTM, and hybrid LSTM-BLSTM models, with ensemble learning for improved accuracy.**	**High Accuracy Feature Selection Optimized Model Training**	**Imbalance in Datasets Dependence on Data Quality**	**The NHIRD dataset includes 90,000 patients, with 18,000 having CKD, featuring 965 comorbidities and 537 medications, collected over two years for health risk prediction**

CONCLUSION :

SMART HEALTH ADVISOR is an innovative AI system aimed at early disease detection, including heart disease, diabetes, and kidney conditions. It uses machine learning models like Random Forest and Logistic Regression to analyze health data such as blood sugar, cholesterol, and blood pressure. With a simple Flask-based interface, it provides personalized preventive care suggestions, focusing on accessibility for remote and underserved areas. Future plans involve wearable integration, enhanced predictions using deep learning, real-time health monitoring, and mobile app development, all to improve healthcare affordability and equity, reducing gaps and supporting global public health.

REFERRENCES

1. Rehman, A., Singh, S., Singh, V., & Hazela, B. (2024). Multiple Disease Prediction System using Machine Learning. Journal of Informatics Electrical and Electronics Engineering (JIEEE), 1-13.

2. Din, F., Haq, A. U., Yaseen, M., Khan, A., & Ali, A. (2024). MULTI-DISEASE PREDICTION USING MACHINE LEARNING AND DEEP LEARNING MODELS. Kashf Journal of Multidisciplinary Research, 1(12), 301-312.

3. Gupta, A., & Singh, A. (2022). An optimal multi-disease prediction framework using hybrid machine learning techniques: 10.48129/kjs. splml. 19321. Kuwait Journal of Science

4. Pradhan, A., Krishna, J. S. V., Kumar, B. P., Tabita, G., Lsastry, V. V. R., & Sudhakar, K. (2024). Integrated Disease Forecasting: Leveraging Deep Learning and Machine Learning for Multi-Disease Prediction. Library of Progress-Library Science, Information Technology & Computer, 44(3).

5. Gaur, N., Fatima, S., SIngh, R. P., & Patel, H. (2024). Multi Disease Prediction System Using Machine Learning.

6. Reshma, K., Niharika, P., Haneesha, J., Rajavardhan, K., & Swaroop, S. MULTI DISEASE PREDICTION SYSTEM USING MACHINE LEARNING.

7. Prakaash, A. S., Sivakumar, K., Surendiran, B., Jagatheswari, S., & Kalaiarasi, K. (2022). Design and development of modified ensemble learning with weighted rbm features for enhanced multi-disease prediction model. New Generation Computing, 40(4), 1241-1279.

8. Anooj, P. K. (2024). An artificial ora algorithm-based recurrent neural network for multi-disease prediction in healthcare monitoring system. JOURNAL OF ARTIFICIAL INTELLIGENCE, 1(2), 11-18.

9. Dhinakaran, D., Raja, S. E., Thiyagarajan, M., Jasmine, J. J., & Raghavan, P. (2024). Optimizing Disease Prediction with Artificial Intelligence Driven Feature Selection and Attention Networks. arXiv preprint arXiv:2408.03151.

10. Zhang, W., Cheng, W., Fujiwara, K., Evans, R., & Zhu, C. (2024). Predictive Modeling for Hospital Readmissions for Patients with Heart Disease: An updated review from 2012-2023. IEEE Journal of Biomedical and Health Informatics.