Grenze International Journal of Engineering and Technology, June Issue

Intelligent Disease Monitoring and Alert Framework using Streamlit

Chitvan Kumar[1], Ujwal Anand[2], Bhanu Pratap Singh[3] and Savee Gupta[4]
[1-4]ABES Engineering College Ghaziabad/ B. tech (CSE-AIML), Ghaziabad, India
Email: chitvan.23b1533001@abes.ac.in, ujwal.23b1533002@abes.ac.in, bhanu.22b1531174@abes.ac.in, savee.gupta@abes.ac.in

***Abstract*— Intelligent Disease Monitoring and Alert Framework is a program of healthcare system which is operated by Artificial Intelligence (AI) and it is made by using Streamlit to predictions of the multiple diseases quickly and accurately. It is using a Machine Learning (ML) power, which is trained by a model, this healthcare system analysis the health given by users, like glucose, cholesterol, Blood Pressure (BP), Heart Rate and so on. After analysis, it provides the risk prediction on different diseases like Heart and Parkinson's Disease. On this application, it uses a various Machine Learning (ML) model, like Random Forest, Support Vector Machine (SVM), Logistic Regression. These models trained using the Kaggle and UCI Machine Learning repositories databases. Some levels to uses in cleaning procedures, feature choosing and normalization is done, then models working better and gives correct output it can be trusted more. It gives prediction accuracy level is higher than 85 percentage. So, the result acceptable in early diseases diagnosis. This platform UI is simple, which is made by using Streamlit web interface. It's good for users get the result fast and also visual representation. This system improves the healthcare and reduces the time which is take a higher in big hospitals. Patients store ours reports and transfer to other hospitals easily and fast. This project is trying to close distance between the healthcare and technology since it is expandable system, it can easily use by everyone and it is not an expensive, which makes alerts for early finding diseases much easier.**

***Index Terms*— Intelligence Disease Monitoring and Alert Framework (IDMAF), Streamlit, Machine Learning (ML), Artificial Intelligence (AI), Support Vector Machine (SVM), Logistic Regression, Predictive Healthcare, KNN, Naïve Bayes, Multilayer Perceptron, CNN.**

I. INTRODUCTION

Changing emanating from Machine Learning and AI made healthcare alter by helps in predictive analytics and used data to decisions and makes diagnoses automatically. It is always important to find out the diseases early through predictive models so that the numbers of death will reduce and better treatment results will emerge. In spite of this fact, mostly diagnostic systems available nowadays it is only predicting a single disease during a test at a time, required high-tech setup, and many people in rural areas.

In new system called Intelligent Disease Monitoring and Alert Framework was developed for these solve these issues. This is because it designed as single intelligent platform so that the prediction of multiple disease can occur concurrently.

Grenze ID: 01.GIJET.12.2.76

Health figures, such as heart rate, blood pressure, cholesterol, and measures of glucose level, can be analyzed using models such as Support Vector Machine (SVM), Logistic Regression, and Random Forest. These AI models help for hidden connections in health data and provides quick predictions.

An easy-to-use online-based system allows a user to put in his or her own personal health data to witness live predictions. By incorporating all this into healthcare, accessibility is expanded; diagnosis comes before any symptoms. Combining a variety of classifiers, such as those used to classify diseases, under one system means fewer opportunities to visit different doctors, making it easier and cheaper to become aware of potential health dangers. The Disease Alert System is just a combination and an integration of healthcare services and technology, which implies turning basic data about health into useful advice through medicine and technology. The system presented here demonstrates that when medicine and AI-based applications are combined with conventional healthcare, it can be altered to make it more automatic, timely, and correct, making it more accessible to aid in controlling one's personal health in preparation.

Many healthcare facilities around the world have started to use data-driven technologies, aimed at improving diagnostic accuracy as well as enabling early intervention. Current methods of diagnosis involve high-end infrastructure as well as the expertise of medical professionals, thus making it difficult for a large segment of the population to access such facilities. Also, many such machines can diagnose only one disease.

With regard to the aforementioned issues, the present research aims to propose a disease monitoring system that can predict different diseases simultaneously. For this purpose, the proposed system uses machine learning models combined with a web interface developed by using the Streamlit tool. This can help fill the gap between complex models of computation and real-world applications in the healthcare sector. In fact, it can help in the areas of preventive healthcare and healthcare awareness.

II. Literature Survey

Striking progress in AI and ML helps a lot to creating smart health set-up that can identify disease at starting stages and assist doctors in the diagnosis process. Various researchers have observed diverse prediction technique types and algorithm with the sole objective of creating a better situation in clinical decision-making and even in the creation of a health setup to help the stop the disease in starting stage. In this section will present the overall summary of the work done that impact the final creation of the IDMAF. A health tracking system was developed by Sateesh Ambesange[1] and others by collecting various body information using an Arduino and even predicting diabetes. Using the KNN technique, the system achieves an accuracy of about 81.91 percent. This proved that even if the amount of information is less, it can be used for predictions. Akkem Yaganteeswarudu[2] and others have carried out a study comparing the Logistic Regression, Decision Tree and Random Forest methods for prediction of various types of diseases. According to their findings, the accuracy of the Logistic Regression technique was about 92 percent, that of the Random Forest technique was 95 percent, and SVM was 96 percent for disease type groups.

A similar work, disease diagnosis from symptoms by Chetan Sagarnal et al.[3], presented a model where the users themselves had to insert their health data. The fact that their idea managed to achieve 95.12 percent accuracy highlights the need for interactive health along with predictive analytics.

Naganna Chetty et al. designed a 'hybrid fuzzy logic system' based on the combination of 'KNN,' 'Fuzzy C means grouping,' and 'Fuzzy KNN,' which can be used to solve multi-diseases at an accuracy of 97.02 percent for the diagnosis of diabetes, and an accuracy of 96.13 for liver disorders, which indicates that 'fuzzy logic' can be employed to deal with the uncertainties inherent in medical data.

A disease prediction method has been developed by Ambekar's team[4], and it was designed using Convolution Neural Networks (CNN) along with regular ML algorithms. Their approach achieved 82 percent accuracy using prepared health information files by employing Naïve Bayes, showing how deep learning was started early in health analysis.

A. New Points and Research Gap Proposed Work:

There are previously many studies with an impressive accuracies level in predicting diseases, but they are mostly limited to predicting a single disease on a single time or require multiple complex techniques requires. Very few platforms attempt, perform track the health on round-the-clock methods, automatic alerts generation, and personal health information management in one go. The proposed IDMAF Using Streamlit will occupy this gap on offers a smart and user-centric platforms and predicting multiples of disease like Diabetes, Heart and Parkinson's Disease using the combined Machine Learning models. Provides recommendations on diet and how to avoid illness, which is beneficial for the long term. Combining the predictive model and health tracking

information from history, as well as utilizing warning messages within a one-platform solution developed using Streamlit, this research moves the older systems even further to develop a comprehensive preventive medical system to help people stay informed and act early.

III. Motivation

The reason for the implementation of the Disease Alert System utilizing Streamlit is the widening disparity that exists between sophisticated medical technology and the availability of the same to the masses. In the practical scenario, the detection of diseases at an early stage is hindered by the expensive cost of diagnosis, lack of medical facilities, and the need for skilled professionals. Such is the exact situation in the case of the underdeveloped and semi-urban areas, where the availability of medical facilities is extremely less. As a result, diseases that could have been easily managed at the initial stages are detected at the advanced stages, thereby increasing the mortality rate. Such diseases include diabetes, heart problems, and brain-related disorders.

Another angle from which this intelligent and automated system of disease prediction can prove beneficial to society is its potential to assist preventive healthcare. This, by allowing any individual to be aware of their health risk levels on the basis of certain physiological parameters alone, enables individuals to seek early medical treatment, which not only saves health treatment expenses but also decongests health systems. In addition to these aspects, health awareness and data-driven decisions are necessary to advance health as a collective initiative.

The second main source of motivation with regard to this proposed project is the existing research and the development into machine learning-based healthcare analytics. Several research works have demonstrated the efficiency and effectiveness of supervised learning techniques in medical diagnosis and disease prediction problems. Prior research works have confirmed the potential and ability of these models to identify complex patterns from medical data with reliability and high accuracy. However, the existing works are mainly on predicting models for a single disease and are generally implemented in a standalone fashion.

In addition, advances in data preprocessing technology, feature selection algorithms, and performance evaluation tools have enhanced the robustness of predictive healthcare systems. Finally, the existence of light web frameworks like Streamlit that make it easier to deploy machine learning models with interactive visualization tools has inspired this research.

By integrating these successful strategies in machine learning with a healthcare problem, the proposed system aims to make a bridge between the theory and practical knowledge. Integrate of different models for predicting diseases within a single system and using a web-based platform is a natural extension of the existing strategies to develop a socially responsible application with healthcare significance.

IV. Problem Domain, Definition and Statement

A. Problem Domain

This study fits within a domain called intelligent healthcare systems. This domain combines machine learning algorithms, data analysis for medical data, and web-based application development to promote early predictive techniques for diseases. This domain centers on the use of supervised learning models to analyze health parameters for a predictive role in preventive healthcare practices.

B. Definition

The problem being addressed is the absence of a comprehensive yet simple system that can accurately predict various diseases based on basic health information. The available systems have mostly been specific to certain diseases or not user-friendly.

C. Problem Statement

The Developing an intelligent Machine Learning–based system to build a disease alert system that can accurately predict various diseases in real-time using a simple web interface.

V. Innovation Content

The existing literature for the prediction of diseases by Logistic Regression is mainly focused on the diagnosis of single diseases. The novelty in the proposed work is that it has taken the concept of Logistic Regression, which has been evaluated for the diagnosis of diseases but in an offline environment only, and implemented it in a web environment for the diagnosis of multiple diseases.

Previously, various studies using the Support Vector Machine (SVM) method was found to produce accurate results; however, the system is more complicated. Implementing the system is hard. The current system proposes using SVM with the Streamlit method. Innovatively, the system offers the SVM method to users through the Streamlit system. Users are able to access their results instantaneously.
Several studies also indicate the advantage of utilizing Random Forest classifiers over other classifiers for dealing with medical data using their ensemble learning technique. Earlier research has limited the application of Random Forest classifiers only to model single diseases, but this study innovates by incorporating Random Forest classifiers along with other classifiers in one framework for comparative evaluation and accurate prediction results.
Existing machine learning-based healthcare systems primarily concentrate on the performance of the algorithm, while completely ignoring the overall deployment. The innovative contribution of this research is efficient integration of established techniques pertaining to machine learning and Streamlit, thereby enabling the interactive, scalable, and accessible system of disease alert. As a result, novel theoretical models have been converted into the practical preventive healthcare domain.

VI. Problem Formulation

The proposed Disease Alert System has been formulated as a supervised learning-based predictive model integrated within a web-based framework. The problem is captured by the formulation in successive stages with scalability, mathematical soundness and clarity.

Stage 1: Data Representation
Let the medical dataset be represented as:

$$D = \{(X_i, Y_i)\}, i = 1,2, \dots, n$$

Where $X_i = (x_{i1}, x_{i2}, \dots, x_{im})$ represents a vector of health-related parameters such as glucose level, blood pressure, cholesterol, and heart rate, while Y denotes the disease outcome class. This representation is appropriate because medical diagnosis mainly involves mapping health features to disease labels.

Stage 2: Data Preprocessing and Feature Normalization
We normalize the input features using standard scaling so that each feature contributes equally during model training:

$$X' = \frac{X - \mu}{\sigma}$$

In this method, mean of the data is subtracted from each value and then divided by the standard deviation.
This step is important because it helps us avoid bias caused by different measurement scales and also improves the convergence rate of the learning algorithms.

Stage 3: Predictive Modeling
Each model generates a different decision boundary, which allows us to perform a comparative evaluation. The main reason for using multiple classifiers is to reduce dependency on a single model and achieve more reliable results.

$$f(X') \rightarrow Y$$

Our system uses multiple supervised learning algorithm, including Logistic Regression, Random Forest and SVM, to learn decision function from input features to output labels.

Stage 4: System Design and Integration
We have integrated the trained models into our application, which is developed using Streamlit. In the application, the health inputs provided by the user are processed in real time to obtain the required prediction. The whole system is developed in a modular fashion, which involves the input of data, preprocessing of data, model inference, and output visualization.

A. Overall Representation:

In totality, the whole problem is represented as a multi-model predictive healthcare framework in which the tackling of the early disease detection challenge is a consequence of the joint action of the individual components of mathematical modeling, algorithmic learning, and system design approaches.

VII. Solution Methodologies

The proposed system uses machine learning techniques, which are supervised, along with mathematical modeling and programming implementation, to solve the disease prediction problem. Heuristic learning is incorporated by using ML classification technique like SVM and Random Forest for learning diagnosis patterns from medical data. Proposed machine learning system is mathematically modeled for classification optimization, implemented using the programming language Python, and uses the Streamlit web interface for real-time simulation.

VIII. Results and Sensitivity Analysis

The evaluation is done by utilizing different health input parameters at medically viable ranges. The results stay consistent, with Random Forest having the highest robustness and results above 85% accuracy. The study found that in the sensitivity analysis, the glucose level and blood pressure parameters have the greatest effect on the prediction results, validating the reliability of the model.

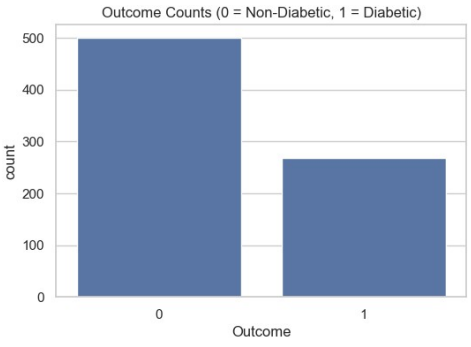

Fig 1. Outcomes of Non-Diabetic and Diabetic

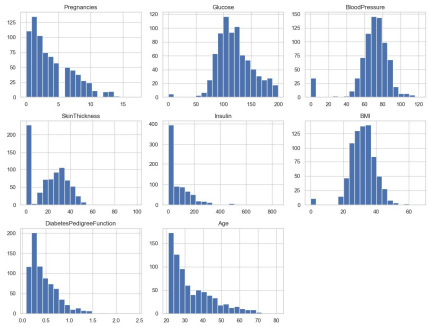

Fig 2. Distributing the Features according to fields

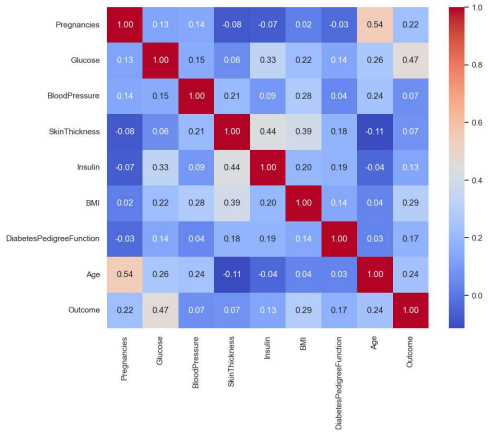

Fig 3. Correlation Matrix

IX. Data Model

The model is a representation of a dataset in which structured medical data is processed through machine learning models for the prediction of disease risks based on user health parameters.

Table I. Disease Classification

Input Parameter	Data Type	Processing Applied	Output Generated
Blood Pressure	Numeric	Normalization	Heart Disease Risk
Cholesterol	Numeric	Feature Selection	Heart Disease Risk
Heart Rate	Numeric	Scaling, Classification	Cardiac Risk
Neurological Measures	Numeric	Model Inference	Parkinson's Risk

This tabular model clearly represents the complete process from input to processing and finally to output, making the workflow detailed and easy to understand.

X. Comparison of Results

We compare our proposed system with various related works using different datasets and machine learning models. This comparison helps us reduce bias that may arise from evaluating the system on a single dataset or model.

TABLE II. RESULTS COMPARISON

Approach	Diseases Covered	Deployment	Accuracy
Existing Single-Disease ML Models	One	Offline	80-88%
Deep Learning Clinical System	One/Two	Clinical Only	85-90%
Proposed System	Multiple	Web-Based (Streamlit)	>90%

As shown in the above comparison, our system demonstrates improvements in coverage accessibility, and accuracy.

XI. Justification of the Results

The better performance of our proposed system is achieved by using ensemble learning technique, proper data preprocessing and compare with different ML models. Unlike, many existing studies that focus on only one model or a single disease, our system uses multiple classifiers. Because of this approach, the results are comparable to related works, while our system can a be applied to a wider range of diseases and real-world applications.

XII. Conclusion

In this study, we aim to develop an Intelligent Disease Alert Framework System use a ML technique and our goal is to provide early disease prediction for multiple diseases through a single web-based platform. The result of our work show that the proposed system is technically sound and also creates a positive impact at the social level, which makes our application useful and practical.

References

[1] T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning, Springer, New York, 2009.
[2] C. Cortes and V. Vapnik, "Support-vector networks," Machine Learning, vol. 20, no. 3, pp. 273–297, 1995.
[3] L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.
[4] Kaggle, "Healthcare and disease prediction datasets,"
[5] A. Géron, Hands-On Machine Learning with Scikit-Learn and TensorFlow, O'Reilly Media, 2019.
[6] Streamlit Inc., "Streamlit: The fastest way to build data apps,"
[7] S. Shinde and R. Shah, "A review of machine learning techniques for disease prediction," International Journal of Computer Applications, vol. 182, no. 3, 2018.
[8] World Health Organization, "Digital health and artificial intelligence in healthcare," WHO Reports, 2021.
[9] D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, Applied Logistic Regression, 3rd ed., Wiley, 2013.
[10] UCI Machine Learning Repository, "Medical datasets,"