EasyChair Preprint

№ 11062

A Symptom Analysis Model for the Detection of Multiple Disease Using Ensemble Machine Learning Algorithm

S Veena, Aniket Kumar, Santhosh R Kumar and Aditya Swaroop

October 9, 2023

A SYMPTOM ANALYSIS MODEL FOR THE DETECTION OF MULTIPLE DISEASE USING ENSEMBLE MACHINE LEARNING ALGORITHM

S. Veena[1] , Aniket Kumar[2] , Santhosh Kumar R[2] , Aditya Swaroop[3]

[1]Associate Professor, Department of Computer Science Engineering, SRM Institute of Science and Technology, Chennai
[2]Department of Computer Science Engineering, SRM Institute of Science and Technology, Chennai
[3]Department of Information Technology, SRM Institute of Science and Technology, Chennai

----------***----------

Abstract:
The primary goal of this application is to identify five unique diseases based on symptom analysis. This application will be based on R programming language, Python, and React, with project data stored on the cloud-based Railway database server. The current illness detection method has various flaws, including low accuracy, long processing times, and exorbitant prices. This project aims to overcome these challenges by analysing symptom data and detecting diseases with high accuracy and speed using powerful machine learning techniques.

The suggested method outperforms the existing approach in various ways, including faster disease diagnosis, higher accuracy rates, cheaper costs, and improved accessibility. Because the system is hosted on a cloud server, users can access it from any location. Furthermore, the system is user-friendly, with an intuitive interface that allows users to readily input their symptoms.

Key Words: Symptom Analysis, Python, React, Detecting Disease, Accuracy

1. INTRODUCTION

This Symptom Analysis Application is a system for detection and prediction the occurrence and progression of diseases in individuals using various data sources such as patient data, genetic data and other health-related information. It is a project which is designed to track and detect five different diseases using advanced machine learning algorithms. The primary purpose of this project is to put together a trustworthy practical solution that can help those who are determining and managing symptoms early, improving their overall health. The application is designed to collect data from users, including their vital signs and self-reported symptoms. This data is then analysed using advanced machine learning algorithms that can identify patterns and trends that may indicate one of the five diseases being tracked.

One of the most important features of this application is its user-friendly interface, which makes it easy for anyone to enter their data and get real-time feedback on their health status. These advanced calculations have the capability to perceive complicated designs and patterns inside the information, empowering the framework to identify the nearness of any of the five infections beneath examination. This explanatory ability marks a critical progression in illness location, promising a more exact and dependable implies of recognizing wellbeing issues based on side effect information.

The project provides a fresh and unique method to disease diagnosis by utilizing advanced machine learning algorithms, cloud-based datasets, and user-friendly interfaces. The application, with its advanced algorithms, user-friendly interface, and inexpensive prices, has the potential to transform disease identification and improve healthcare outcomes for people all over the world.

The five diseases which can be detected using this application are:

1. Mesothelioma

Mesothelioma is a cancer from developing of the epithelium of the respiratory tract, the gastrointestinal tract, or heart. It can often be triggered by ingesting asbestos and can take

years to come about. Symptoms include chest pain, shortness of breath, and fatigue, and treatment may include surgery, chemotherapy, or radiation.

2. Coronary Heart Disease

Coronary Heart Disease is a common cardiovascular medical condition in which the blood coronary arteries that hydrate the heart shrink or become obstructions owing to accumulation of plaque. It has the ability for triggering difficulty breathing, loss of respiration, and a cardiac event. Treatment may include lifestyle changes, medications, or procedures such as angioplasty or bypass surgery.

3. Diabetes Mellitus

Diabetes Mellitus is an autoimmune disease that alters the manner in which the body absorbs sugar in the blood (glucose). Type 1 diabetes is an aggressive scenario that requires injections of insulin, whereas Type 2 diabetes is quicker to develop and is most frequently dictated with physical activity, diet, and prescription.

4. Liver Cirrhosis

Liver Cirrhosis is a condition in which the liver becomes damaged and scarred over time. This can be caused by alcohol abuse, hepatitis or other diseases. Symptoms include fatigue, weakness, and jaundice, and treatment may include lifestyle changes, medications, or a liver transplant.

5. Chronic Renal Disease

Chronic Renal Disease is a medical condition which prompts the renal system to lose their functionality as time goes on. Diabetes may have a role in this, high blood pressure or other factors. Symptoms may not appear until the disease is advanced, and treatment may include medications, dialysis, or a kidney transplant.

2. LITERATURE REVIEW

Maad M. Mijwil, Israa Ezzat Salem and Rana A. Abttan [1] stated about chemical waste, hazardous contamination, and everyday stresses all contribute to disorders like COVID-19. In order to discover the strategy that is most significant and which is more serious, he additionally contrasted how each method was put into practice.

Maad M. Mijwil and Ban Salman Shukur [2] showed the Machine Learning methodologies that play a crucial and significant part in interpreting an illness information, anticipating coronary artery disease, and supporting deciding.

Ibrahim Mahmood Ibrahim, Adnan Mohsin Abdulazeez [3] has provided an overview machine learning algorithms used in medical diagnosis, including Naïve Bayes, logistic regression, support vector machine, K-nearest neighbour, comparing their effectiveness.

R. L. Priya, S. Vinila Jinny [4] proposed from the paper that in order to identify persistent illnesses in older individuals early and minimize catastrophes, multiple machine learning models for prediction were investigated incorporated with, or without, other connected analytical strategies. The pros and drawbacks of different models of prediction have been looked at as well.

Christopher Toh along with James P. Brody [5] stated the numerous formats in which medical information can be found, also presented some issues in data structuring and can increase noise. They also talked about a brief history of machine learning and the present state of this technology in healthcare.

Rayan Alanaz [6] In his paper, he addressed the system that had been suggested, that employs algorithmic methods for machine learning for feature extraction and forecasting diseases for determining the distance in order to offer expansive prediction of disease based on symptoms reported by patients.

Jian Huang [7] used image analysis technology built around machine learning has been employed in the identification of paediatric diseases, and a future application of technological advances in image analysing to the surveillance of infantile health conditions is explored. It also alleges that Machine Learning technology is capable of identifying white blood cells that are tricky to see with one's own eyes.

Feng Hao, Kai Zheng [8] For aiding clients in receiving consultations via the internet, a model

has been developed by bringing together machine learning techniques with data graph advances in technology. The tool uses frequent appearances analysis of words for evaluating classification data from various divisions. The inquiry into the topic revealed that the machine learning method framework was 96.29% effective for recognizing things in computerized healthcare information.

Rijwan Khan [9] explored the impact of machine learning on healthcare and the creation of an application called "Medicolite" that includes multiple modules for assisting with health-related difficulties such as dietary concerns. It also offered online doctor appointments and medicines over the phone.

R.P, Prawin & R.P, Pranav [10] noted that the implementation of machine learning techniques for renal failure detection which can be essential for the fast detection and evaluation of kidney illness. They compared different types of model and finds the best approach to predict the illness.

Md Manjurul Ahsan , Shahana Akter Luna and Zahed Siddique [11] had summarised the comprehensive study of trending methods in machine learning based disease diagnosis and also included the use of different types of algorithms, different types of disease and evaluation metrics.

Talasila Bhanuteja [12] created an illness prediction scheme via machine learning estimations such as a random forest technique and the decision-tree serves proficiently and explicates the negligent application of information mining methods in industry, management of medicine, and scientific research that have promoted the implementation of Machine Learning (ML) in clinical data sets for early disease prediction, patient care, and local administration.

Aseel Alfaidi [13] provides a framework which relies on machine learning (ML) structures to forecast someone's probabilities to succumb to cardiomyopathy, with the plurality of layers of delivering the most accurate cancer forecasting precision with 87.23%.

Balajee R M [14] focused on the discussion of medical needs of researchers while dealing with automated prediction application. They additionally addressed the optimal predictive modelling strategy to recognize healthcare professionals records.

Sankavi Muralitharan [15] proposed a model which evaluates machine learning-based early warning systems for predicting physiological deterioration in acutely ill patients. It came across that such frameworks are significantly more precise over aggregate-weighted systems as a whole but there still exists plenty of room for enhancement.

3. METHODOLOGY

3.1. DATASET

The dataset from Kaggle for my project is an important part of our mission to detect five different diseases through symptom analysis. Kaggle's platform provides many important data resources for training and validating our machine learning algorithms.

With algorithm selection, hyperparameter tuning, and thorough model evaluation, your project aims to deliver reliable and impactful results. Ethical considerations, such as addressing biases and safeguarding patient privacy, are integral to responsible AI development, particularly in healthcare.

Attributes	Type
Age	Numbers
Gender	Choosing Values
Smoking	Choosing Values
Yellow Fingers	Choosing Values
Anxiety	Choosing Values
Peer Pressure	Choosing Values
Chronic Disease	Choosing Values
Fatigue	Choosing Values
Allergy	Choosing Values
Wheezing	Choosing Values
Alcohol	Choosing Values
Cough	Choosing Values
Shortness of Breathe	Choosing Values
Swallowing Difficulty	Choosing Values
Chest Pain	Choosing Values

Table 1. List of Attributes for Mesothelioma

Attributes	Type
Age	Numbers
Gender	Choosing Values
Chest Pain Type	Choosing Values
Resting Blood Pressure	Numbers
Serum Cholesterol	Numbers
Is Fasting Blood Sugar is greater than 120 mg/dl?	Choosing Values
Resting Electrographic Results	Choosing Values
Maximum Heart Rate Achieved	Numbers
Exercise Induced Angina	Choosing Values
Old Peak	Numbers
Slope of the Peak Exercise	Choosing Values
Number of the Major Vessels	Choosing Values

Table 2. List of Attributes for Coronary Heart Disease

Attributes	Type
Age	Numbers
No. of Pregnancies	Choosing Values
Glucose Level	Choosing Values
Blood Pressure	Numbers
Skin Thickness	Numbers
Insulin	Choosing Values
BMI	Choosing Values
Diabetes	Numbers

Table 3. List of Attributes for Diabetes Mellitus

Attributes	Type
Age	Numbers
Gender	Choosing Values
Presence of Ascites	Choosing Values
Presence Of Hepatomegaly	Choosing Values
Presence of Spiders	Choosing Values
Presence of Edema	Choosing Values
Serum Bilirubin	Numbers
Serum Cholesterol	Numbers
Allergy	Numbers
Albumin	Numbers
Urine Copper	Numbers
Alkaline Phosphate	Numbers
SGOT	Numbers
Triglycerides	Numbers
Platelets	Numbers
Prothrombin	Numbers

Table 4. List of Attributes for Liver Cirrhosis

Attributes	Type
Age	Numbers
Blood Pressure	Numbers
Resting Blood Pressure	Numbers
RBC in Urine	Choosing Values
POS Cell in Urine	Choosing Values
POS Cell Clumps in Urine	Choosing Values
Bacteria in Urine	Choosing Values
Blood Glucose Random	Numbers
Blood Urea	Numbers
Serum Creatinine	Numbers
Sodium	Numbers
Potassium	Numbers
Haemoglobin	Numbers
Packed Cell Volume	Numbers
WBC Count	Numbers
RBC Count	Numbers
Hypertension	Choosing Values
Diabetes Mellitus	Choosing Values
Coronary Artery Disease	Choosing Values
Appetite	Choosing Values
Pedal Edema	Choosing Values
Anaemia	Choosing Values

Table 5. List of Attributes for Chronic Kidney Disease

3.2 DATA PREPARATION

The overarching objective of data preparation includes to turn raw information into a format that machine learning algorithms can understand. Pre-processing helps to improves the accuracy and usefulness of algorithms for machine learning by utilizing numerous approaches and procedures. As a result, these algorithms are better able to detect patterns and

relationships generating the information, thereby resulting in more precise and meaningful predictions and insights.

1) Data Cleaning: There are instances of missing values indicated as question marks inside the collection of data that must be replaced with some values. Furthermore, the dataset may contain incorrect data, that is capable of being identified via data profiling and statistical analyses.

2) Managing Missing Values: For keeping track of the value gaps throughout this the collection of data, one can employ anticipate addresses including maximum replacement for disproportionate quantitative information and mode replacement for classification data.

3) Detecting Outliers: The outliers are records that diverge enormously beyond the remainder of the collected information and may decrease the precision of the machine learning strategies. As a result, detecting and addressing outliers in the dataset is critical, which can be achieved using methods such as Boxplots, Z-score, or Interquartile Range.

3.3 ARCHITECTURE DIAGRAM

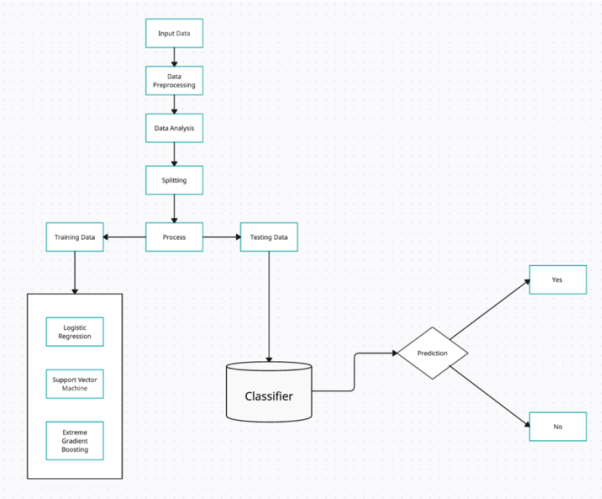

Fig: Architecture Diagram of the proposed system

3.4. ALGORITHMS IMPLEMENTED

3.4.1 Logistic Regression

It is a supervised learning technique used in machine learning. When determining if something is possible or not (0/1, Yes or No), it is utilized to forecast the binary outcomes. It forecasts the likelihood of a dependent variable using the logistic function. In order to produce the predictions, a threshold of (0.5) was selected; if an instance is above the threshold, a positive class (1) will be created, otherwise 0.

In our application, logistic regression can be used as an initial approach for binary classification tasks like determining whether a specific disease is present 1 or absent 0. The first step in using logistic regression is to prepare the dataset by categorizing the symptoms as features and the disease's presence or absence as the target variable.

Logistic Regression is particularly well suited to binary classification challenges since it models the likelihood of an occurrence belonging to a specific class. It is possible to avoid overfitting by using regularization techniques like L1 or L2 regularization. The performance of the model would need to be enhanced through careful tuning of the hyperparameters like the regularization strength. In the end, Logistic Regression may provide an intuitive yet efficient baseline model for disease identification in the project.

3.4.2 Support Vector Machine

It is an approach of supervised learning that tends to be used for tasks involving classification works and regression factor. The overarching objective of SVM is to seek a shape known as a hyperplane that splits the information provided into two distinct sets. The Support Vector Machine finds optimal hyperplane that boosts the margin between classes and delivers accurate predictions for both classification and regression problems. The system's performance and ability to handle different types of data are significantly influenced by the kernel function and hyperparameters that are chosen.

Support Vector Machines (SVM) can be an effective tool for this categorization of disease. SVMs are very adept at binary and multiclass classification tasks. Scaling the dataset's features and encoding the target variable are necessary steps in the pre-processing phase of SVM implementation. Then, to transfer the data into a higher-dimensional space, an appropriate

kernel function, such as a Linear or Radial Basis Function (RBF), would be used.
To maximize SVM performance, hyperparameters such as the kernel parameter and regularization parameter should be fine-tuned using approaches such as grid search or cross-validation. SVM's ability to locate the best hyperplane for classification makes it particularly beneficial for complicated illness detection applications with non-linear decision boundaries.

3.4.3 Extreme Boosting

Extreme boosting, often known as XGBoost, is a strong ensemble machine learning approach known for its speed and accuracy. It combines the advantages of decision trees with gradient boosting and uses methods like regularization and parallel processing to swiftly handle large datasets.

XGBoost, an ensemble learning method, could significantly enhance disease detection accuracy for this project. To use XGBoost, start by preprocessing the dataset as previously. XGBoost works by constantly correcting faults caused by earlier trees while training an array of decision trees. To find a balance between model complexity and accuracy in prediction, hyper parameters like as learning rate, maximum tree depth, and number of boosting rounds should be fine-tuned.

XGBoost's performance may also be enhanced via enhancing feature design and selection. XGBoost is a significant boost to the research due to its capacity to handle unbalanced datasets as well as recognize challenging relations between signs and symptoms, which could contribute to improved diagnosis of diseases precision and dependability.

4. RESULTS AND DISCUSSION

Performance Evaluation: The performance of a prediction model is measured to make certain that it is completely present in the evidence set as well as that it serves adequately on untouched data. The intent of an outcome analysis is to figure out how accurate a model is at generalizing to touched/un-touched information. According to True Positive, True Negative and False Positive, False Negative, the classifier's performance may be characterized, which are used to figure out the performance of different binary classification models. The four measures mentioned above are used to get a better idea of how accurate and reliable a model is. True positives (TP) are positive predictions that turn out to be actually positive. False positives (FP) are positive predictions that turn out to be actually negative. True negatives (TN) are negative predictions that turn out to be actually negative. False negatives (TN) are negative predictions that turn out to be actually positive.

Accuracy: It is the parameter used to evaluate categorization of different types Machine Learning algorithms. The proportion of right assumptions generated via the model we use is referred to as accuracy.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision: It is the proportion of favourable outcomes irrespective of all projected instances that are favourable. Precision answers the question “What proportion of positive identifications was actually correct?”

$$\text{Precision} = \frac{TP}{TP + FP}$$

Recall: It is a proportion of favourable happenings irrespective of all actual positive occurrences. Recall simplifies question “What proportion of actual positives was identified correctly?”

$$\text{Recall} = \frac{TP}{TP + FN}$$

F1-Measure: The F1 score is the harmonic mean of exactness and retention. Since all of them influence this, the superior the F1 score, the more desirable it is. Therefore we observe, if either is low as well, the overall F1 rating diminishes extensively as a result of the resultant value within the fraction.

$$\frac{2}{\frac{1}{precision} + \frac{1}{recall}} = \frac{2 * precision * recall}{precision + recall}$$

Mesothelioma:

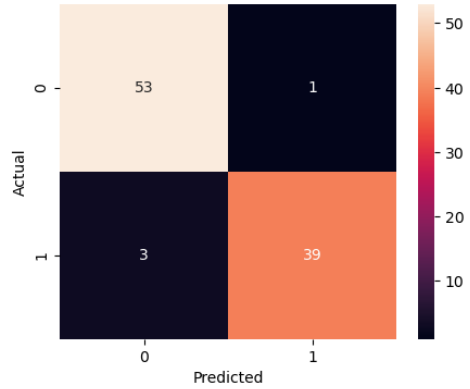

Fig: Matrix of Mesothelioma

	precision	recall	f1-score	support
0	0.95	0.98	0.96	54
1	0.97	0.93	0.95	42
accuracy			0.96	96
macro avg	0.96	0.96	0.96	96
weighted avg	0.96	0.96	0.96	96

Fig: Classification of Mesothelioma

Coronary Heart Disease:

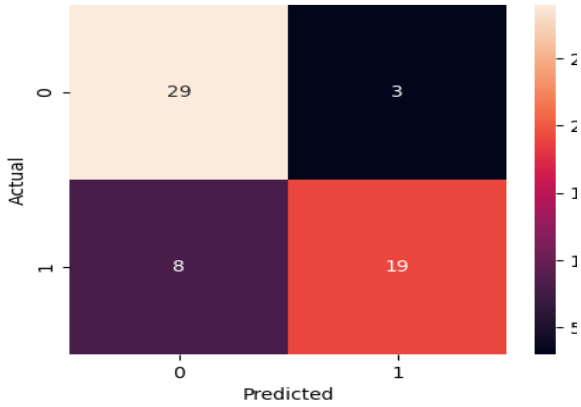

Fig: Matrix of Coronary Heart Disease

	precision	recall	f1-score	support
0	0.78	0.91	0.84	32
1	0.86	0.70	0.78	27
accuracy			0.81	59
macro avg	0.82	0.80	0.81	59
weighted avg	0.82	0.81	0.81	59

Fig: Classification of Coronary Heart Disease

Diabetes Mellitus:

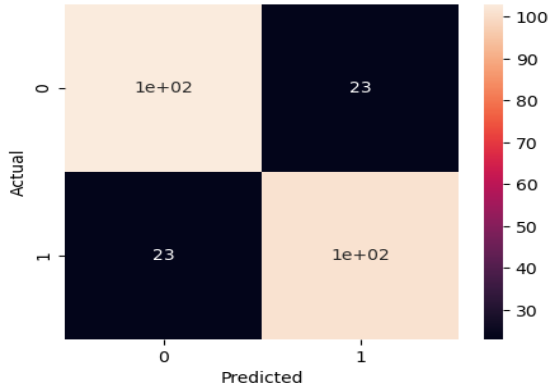

Fig: Matrix of Diabetes Mellitus

	precision	recall	f1-score	support
0	0.82	0.82	0.82	126
1	0.81	0.81	0.81	124
accuracy			0.82	250
macro avg	0.82	0.82	0.82	250
weighted avg	0.82	0.82	0.82	250

Fig: Classification of Diabetes Mellitus

Liver Disease:

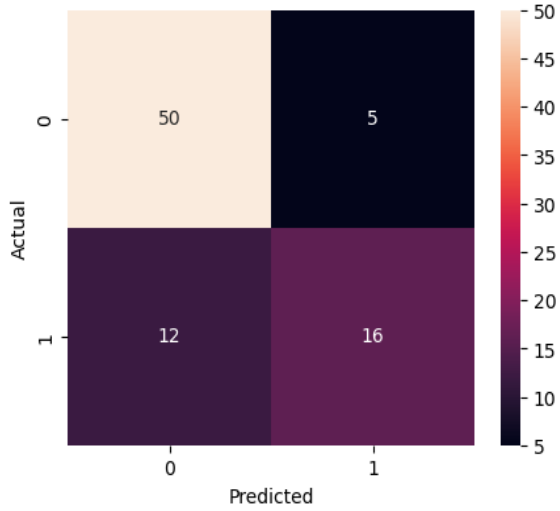

Fig: Matrix of Liver Cirrhosis

	precision	recall	f1-score	support
0.0	0.81	0.91	0.85	55
1.0	0.76	0.57	0.65	28
accuracy			0.80	83
macro avg	0.78	0.74	0.75	83
weighted avg	0.79	0.80	0.79	83

Fig: Classification of Liver Cirrhosis

Chronic Kidney Disease:

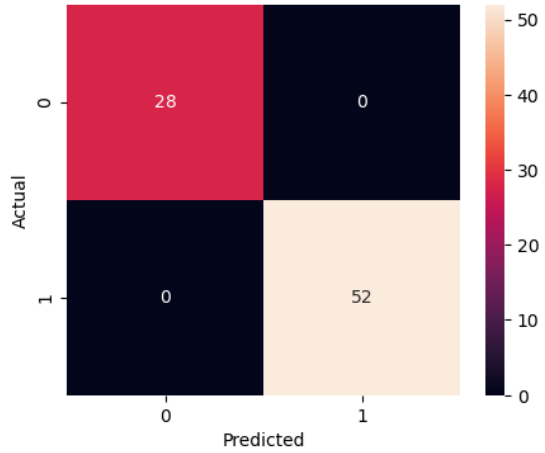

Fig: Matrix of Chronic Kidney Disease

	precision	recall	f1-score	support
0	1.00	1.00	1.00	28
1	1.00	1.00	1.00	52
accuracy			1.00	80
macro avg	1.00	1.00	1.00	80
weighted avg	1.00	1.00	1.00	80

Fig: Classification of Chronic Kidney Disease

5. RESULTS AND DISCUSSION

In summary, by giving people an effective tool to control and prevent chronic diseases, this research significantly improves upon the current healthcare system. The goal of the study was to create a tool that would aid users in managing and preventing chronic diseases by utilizing cutting-edge technology including machine learning algorithms, data analysis, and real-time feedback.

These goals were accomplished by the project successfully, and the resulting technology has the potential to completely change the healthcare sector. The system is well-designed to gather user information about symptoms, vital signs, and lifestyle choices. This information is then analysed using machine learning algorithms to find probable ailments like diabetes, heart disease, lung disease, depression, and hypertension.

The early detection and treatment of many diseases was found to benefit from the proposed approach, improving patient outcomes and lowering complications. The study's findings may serve as a springboard for additional research in this area because they shed light on the best method for the early detection and treatment of renal failure.

6. FUTURE ENHANCEMENTS

There are many potential for this project to develop and improve in the future. The program would become more comprehensive and beneficial to users if the range of detectable disorders were broadened to include a wider array of typical and potentially fatal conditions. The accuracy and depth of disease diagnosis could also be greatly improved by integrating other health data sources, such as genetic data and medical imaging.

The project could develop to provide individualized health insights in addition to symptom analysis, enabling individuals to take charge of their health. The continuous tracking of health would enable early intervention and a comprehensive understanding of health patterns over time. Real-time monitoring for chronic diseases and integration with wearable devices would ensure this. To safeguard consumers' private health information, it is crucial to prioritize strong privacy protections and security measures.

Additionally, creating a welcoming user community through discussion boards or support groups can improve the psychological and educational components of healthcare involvement. In order to validate the system through clinical trials and keep this application at the forefront of healthcare innovation, collaborations with research institutes and healthcare professionals can further improve disease detection algorithms. These enhancements for the future may result in the increase of scope in this medical field.

7. REFERENCES

[1] Maad M. Mijwil, Israa Ezzat Salem and Rana A. Abttan. Utilisation of Machine Learning Techniques in Testing and Training of Different Medical Datasets. Asian Journal of Computer and Information Systems (ISSN: 2321 – 5658) Volume 9 – Issue 4, October 2021

[2] Maad M. Mijwil, Ban Salman Shukur. A Scoping Review of Machine Learning Techniques and Their Utilisation in Predicting

Heart Diseases. Ibn Al-Haitham Journal for Pure and Applied Sciences. Received 1 March 2022, Accepted 8 May 2022, Published in July 2022.

[3] Ibrahim Mahmood Ibrahim, Adnan Mohsin Abdulazeez. The Role of Machine Learning Algorithms for Diagnosing Diseases. Journal of Applied Science and Technology Trends (JASTT). Vol. 02, No. 01, pp. 10 – 19 (2021)

[4] R. L. Priya, S. Vinila Jinny. Elderly Healthcare System for Chronic Ailments using Machine Learning Techniques – a Review. Iraqi Journal of Science, 2021, Vol. 62, No. 9, pp: 3138-3151 DOI: 10.24996/ijs.2021.62.9.29

[5] Christopher Toh and James P. Brody. University of California, Irvine, United States of America. Applications of Machine Learning in Healthcare. IntechOpen . © 2020 The Author(s). Licensee IntechOpen.

[6] Rayan Alanazi. Identification and Prediction of Chronic Diseases Using Machine Learning Approach. Hindawi Journal of Healthcare Engineering Volume 2022, Article ID 2826127, 9 pages

[7] Jian Huang, Jing Li, Zheming Li, Zhu Zhu, Chen Shen, Guoqiang Qi and Gang Yu. Detection of Diseases Using Machine Learning Image Recognition Technology in Artificial Intelligence. Hindawi Computational Intelligence and Neuroscience Volume 2022, Article ID 5658641, 14 pages

[8] Feng Hao and Kai Zheng. Online Disease Identification and Diagnosis and Treatment Based on Machine Learning Technology. Hindawi Journal of Healthcare Engineering Volume 2022, Article ID 6736249, 7 pages

[9] Rijwan Khan, Akhilesh Kumar Srivastava, Mahima Gupta, Pallavi Kumari and Santosh Kumar. Hindawi. Computational Intelligence and Neuroscience Volume 2022, Article ID 8109147, 12 pages

[10] R.P, Prawin & R.P, Pranav & Rajavel, Swathi. (2023). Performance Evaluation and Comparative Analysis of Several Machine Learning Classification Techniques Using a Data-driven Approach in Predicting Renal Failure. International Journal for Research in Applied Science and Engineering Technology. 11. 3522-3530. 10.22214/ijraset.2023.54343.

[11] Md Manjurul Ahsan, Shahana Akter Luna and Zahed Siddique. Machine-Learning-Based Disease Diagnosis: A Comprehensive Review. Healthcare 2022,10,541. Received: 10 February 2022 Accepted: 10 March 2022 Published: 15 March 2022

[12] Talasila Bhanuteja, Kilaru Venkata Narendra Kumar, Kolli Sai Poornachand, Chennupati Ashish, Poonati Anudeep. Symptoms Based Multiple Disease Prediction Model using Machine Learning Approach. International Journal of Innovative Technology and Exploring Engineering (IJITEE)

[13] Aseel Alfaidi, Reem Aljuhani, Bushra Alshehri, Hajer Alwadei, Sahar Sabbeh. Department of Computer Science and Artificial Intelligence, University of Jeddah, Jeddah, KSA1, 2, 3, 4, 5 Faculty of Computer Sciences and Artificial Intelligence, Benha University, Egypt

[14] Balajee R M , Hitesh Mohapatra and Deepak V, Vijendra Babu D. Requirements Identification on Automated Medical Care with Appropriate Machine Learning Techniques. 2021 6th International Conference on Inventive Computation Technologies (ICICT)

[15] Sankavi Muralitharan, MPharm, MSc; Walter Nelson1*, BSc; Shuang Di, BSc, MEd, MSc; Michael McGillion, BScN, PhD; PJ Devereaux, MD, PhD, FRCPC; Neil Grant Barr, BA, MSc, PhD; Jeremy Petch, HBA, MA, PhD. Machine Learning–Based Early Warning Systems for Clinical Deterioration: Systematic Scoping Review. JOURNAL OF MEDICAL INTERNET RESEARCH

[16] Kaustubh Arun Bhavsar, Ahed Abugabah, Jimmy Singla, Ahmad Ali AlZubi, Ali Kashif Bashir and Nikita. A Comprehensive Review on Medical Diagnosis Using Machine Learning. CMC, 2021, vol.67, no.2. Received: 28 October 2020; Accepted: 14 December 2020

[17] Dr. T. Kameswara Rao, P. Chamanthi, N. Tharun Kumar, R. Lakshmi Amulya, M. Uday Sagar. Skin Disease Detection Using Machine Learning. IJFANS INTERNATIONAL

JOURNAL OF FOOD AND NUTRITIONAL SCIENCES. Volume 11, Iss 12, Dec 2022

[18] Maxim E. Zaslavsky, Nikhil Ram-Mohan, Joel M. Guthridge, Joan T. Merrill, Jason D. Goldman, Ji-Yeun Lee, Krishna M. Roskin, Charlotte Cunningham-Rundles, M. Anthony Moody, Barton F. Haynes, Benjamin A. Pinsky, James R. Heath, Judith A. James, Samuel Yang, Catherine A. Blish, Robert Tibshirani, Anshul Kundaje and Scott D. Boyd6. Disease diagnostics using machine learning of immune receptors. April 28, 2022

[19] Firdews A.Alsalman, Shler Farhad Khorshid, Amira Bibo Sallow. Disease Diagnosis Systems Using Machine Learning and Deep learning Techniques Based on TensorFlow Toolkit: A review. Al-Rafidain Journal of Computer Sciences and Mathematics (RJCM), Vol. 16, No. 1, 2022

[20] Madhumita Pal, Smita Parija, Ganapati Panda, Kuldeep Dhama, Ranjan K. Mohapatra. Risk prediction of cardiovascular disease using machine learning classifiers. Received February 11, 2021; accepted May 23, 2022

[21] Rabeea Mansoor , Munam Ali Shah , Hasan Ali Khattak, Shafaq Mussadiq, Hafiz Tayyab Rauf and Zoobia Ameer. Detection of Diseases in Pandemic: A Predictive Approach Using Stack Ensembling on Multi-Modal Imaging Data. Received: 30 October 2022 Accepted: 21 November 2022 Published: 30 November 2022

[22] Mohammad Abood Kadhim, Abdulkareem Merhej Radhi. Heart disease classification using optimized Machine learning algorithms. Received November 2022; Accepted Janury 2023; Available online Februry 2023.

[23] Bayan Alsaaidah, Moh'd Rasoul Al-Hadidi, Heba Al-Nsour, Raja Masadeh and Nael AlZub. Comprehensive Survey of Machine Learning Systems for COVID-19 Detection. Received: 12 July 2022 Accepted: 20 September 2022 Published: 30 September 2022.

[24] Almira Suljovi, Stevan Čakić and Tomo Popović. Detection of Plant Diseases Using Leaf Images and Machine Learning. 21st International Symposium INFOTEH-JAHORINA, 16-18 March 2022

[25] Omar Enzo Santangelo, Vito Gentile, Stefano Pizzo, Domiziana Giordano and Fabrizio Cedrone. Machine Learning and Prediction of Infectious Diseases: A Systematic Review. Received: 7 December 2022 Revised: 20 January 2023 Accepted: 30 January 2023 Published: 1 February 2023