J. Electrical Systems 20-9s (2024): 2119-2129

[1]Shakeel Juman T P

[2]Dr. K. Aravinthan

Improved Health Care Diagnostic and Preventative System Using Artificial Intelligence Approach

Abstract: - The data health organizations generate is so large and complex that it can be difficult to analyze to make important decisions about patient health. These data contain detailed information about hospitals, patients, medical claims, treatment costs, and more. Therefore, there is a need to generate a powerful tool to analyze and extract important information from these complex data. Machine learning (ML) in clinical benefits is one of the most exciting areas of investigation. Machine learning plays an important role in spotting new trends in healthcare organizations, which in turn is useful to others related to the field. It is a subfield of software engineering that uses existing data in different databases to transform it into new research findings. The aim of this research is to help physicians use a tool to quantify health disorders through medical data that can be implemented in a computer-aided system. Computer-aided assessments are always faster than manual processes and provide physicians with additional resources for diagnosis and treatment. The proposed method defines specific patterns of disease in health using machine learning techniques. Compared with other existing algorithms, the proposed Improved Healthcare Prediction Using Artificial Intelligence Method (IHPAI) system yields higher efficiency in predicting healthcare-related diseases.

Keywords: Healthcare, prediction methods, automatic predictions, health, computing aids, prediction guardians, artificial intelligence

I. INTRODUCTION

Data analysis using Data Mining (DM) techniques have been adjusted to many domains in addition to statistical analysis. DM techniques have been used in healthcare, marketing, business, social media and any field having voluminous data. Visualization and interpretation of results using discovered knowledge is the main task of a technique in its analysis [1]. DM or Knowledge discovery is a growing field driven by research interests and needs and an essential area of research in healthcare. Healthcare uses DM techniques for knowledge discovery and identification of successful disease prescribing patterns and predictions using computer-aided diagnosis or expert learning. DM and forecast integration can provide high-quality and reliable forecasts. Sickness expectation utilizing DM strategies is a spurring task to improve diagnostic accuracy. Hence the objective of this research is in using DM as they help decline cost and time. Information disclosure from clinical information is a convoluted undertaking, mostly because of insignificant and undesirable information.

Using more than one DM technique for predicting diseases can also result in better accuracy. Hence the main objective of this research work is to predict diseases from patient's records and suggesting a non-invasive DM model. Moreover, Features provide state-of-the-art performance for recognition of abnormalities. While the accuracy of action recognition has been continuously improved over the recent years, the extraction of lesser number of features and subsequent identifications based on these extractions have been preventing methods from scaling up to real-life issues [2]. This problem is addressed in this research work by the development of highly efficient features using feature information in disease recognitions. Moreover the speed of feature extraction and feature selection can help disease classification perform better at the cost of a negligible reduction in recognition accuracy. The main goal of this work is efficient disease recognition while exploring speed and memory trade-offs in feature extraction and selection.

This work uses a combination of DM techniques for reducing the complexity in historical disease data and identifying disease [3]. The objectives of the research work is detailed below

- Developing a non-invasive model for predicting disease

[1] Ph.D Research Scholar, Department of Computer Science, Adaikalamatha College, affiliated to Bharathidasan University, Vallam, Thanjavur, Tamilnadu, shakeeljuman81@gmail.com

[2]Assistant Professor, Department of Computer Science, Adaikalamatha College, affiliated to Bharathidasan University, Vallam, Thanjavur, Tamilnadu, aravinthk83@gmail.com

- Examine Disease Data and prepare them by pre-processing
- Coin a new Dataset which can be utilized for nonexclusive illness expectation.
- Classify and predict disease from patient's data.
- Improving the accuracy of disease in CAD systems and speed.
- Identify the required minimum set of features for predicting diseases.
- To implement a "Lupus Prediction System" for speed and performance accuracy.

II. LITERATURE SURVEY

Kishor, *et. al.*, [2022] suggested a model that makes use of feature selection strategies to classify the thyroid data. They improved the developed model's performance by utilising feature selection approaches.

Krishnamoorthy, *et al.*, [2023] An automatic multi-layer perception (Auto MLP) application for diabetes prediction is designed. This technique also uses improved class outlier detection. The parameters are automatically adjusted during the training process. Outlier detection is performed during data preprocessing.

Dwivedi, Ruby, *et al.*, [2022] Analyze the big data generated by cloud-based medical IoT devices instead of relying on the limited computing and storage resources of handheld devices; propose a health monitoring and diagnosis framework based on cloud computing under the Internet of Things, which can predict the severity of underlying diseases; by using Diagnostic scenarios for multiple state-of-the-art classification algorithms, which were found to outperform reference methods in terms of accuracy, sensitivity, specificity, and F1 in disease prediction.

Vimont, *et al.*, [2022] the paper reviews artificial intelligence solutions for dementia care in healthcare informatics. It examines and evaluates current logical systems, identifying fundamental issues and challenges with having sufficient healthcare data. It also addresses future possibilities and implications for serious AI research, such as understanding dementia informatics.

Rastogi, *et al.* [2023] showed that effective early liver infection acknowledgment through Multi-facet Insight Brain Organization estimation relies upon various decision tree calculations, such as chi-square programmed communication indicator and characterization, and relapse tree with boosting strategy. Their technique had the option to analyze and characterize the liver malady proficiently.

Abdollahi, *et al.,[*2022] It shows that artificial intelligence has been applied in agriculture and healthcare to improve crop production, disease prediction, continuous monitoring, efficient supply chain management, operational efficiency, and water waste. The main goal is to design standard and reliable product quality control methods, and to find New Approaches Deep learning and machine learning are two of the most popular artificial intelligence techniques.

Ghaffar Nia, *et al*., [2023] is Data mining for healthcare is an interdisciplinary topic of research that evolved from database statistics and is valuable in assessing the efficacy of medical interventions. Data visualization with machine learning Diabetes-related coronary illness is a sort of coronary illness that happens in diabetics.

Kapoor, *et al.*, [2023] the paper reviews artificial intelligence solutions for dementia care in healthcare informatics. It examines and evaluates current logical systems, identifying fundamental issues and challenges with having sufficient healthcare data. It also addresses future possibilities and implications for serious AI research, such as understanding dementia informatics.

III. PROPOSED WORK

Automatic disease identification using computer aided systems has gotten significant consideration over the most recent twenty years. Challenges posed on this subject include symptoms are often not well defined and the symptoms produced by different diseases may overlap or be present simultaneously. The scope of this study was to provide a model to overcome some of the challenges faced by automated disease detection. Lupus is a multisystem disease that commonly affects women [4] .Disease prediction results in manifestation of other symptoms like cutaneous, cardiac, renal, neuropsychiatric, gastro intentional etc. The etiological factor of lupus is multi-factorial as the disease is It is characterized by the production of autoantibodies, leading to the deposition of immune complexes, inflammation and eventually permanent organ damage. Assessment of lupus is not based on single test and it involves accurate physical and laboratory diagnosis.

Recording and monitoring of disease activity for the patient’s improved health status and quality of life has been a major area of research. Lupus is characterized by periods of relative quiescence and periods of exacerbations which involve any organ or system in various combinations of body have to be analyzed.[5] . Appropriate management of lupus is critically dependent Legitimate evaluation of infection movement and personal satisfactions. Hence, the next main aim of this research work is to utilize the knowledge of previous history on disease effectively. Further, discovering hidden patterns and relationship in disease data has not been exploited so often. Hence, the next objective lies in utilizing existing data on disease and creating a model for predicting disease in patients undiagnosed as disease using DM techniques. Thus, the aims of this work revolve around predicting disease based on their symptoms. The main objectives of the research are -

- To extend the life span of a chronically diseased patient.
- Automatically predict a diseased patient from recorded symptoms
- To propose an intelligent clinical decision making model for reducing errors in disease decisions.
- To extend the lifetime of chronically diseased patients and assist with working on the nature of treatment in chronic diseases.
- To design and propose new algorithms and techniques to predict diseases early.
- To study and analyze lupus patients real-time data set and turn it into useful knowledge.
- To propose a new algorithm to predict disease early.
- To increase the speed of predictions while reducing the cost of automated predictions by selecting fewer and important features
- To analyze existing algorithms and compare the proposed method’s effectiveness.

IV. METHODOLOGY

4.1 Dataset

4.1.1 Systemic lupus erythematosus

Systemic lupus erythematosus (SLE) (MIM no. 152700), a chronic systemic inflammatory disease, is considered to be the prototypic example of systemic autoimmune disease. The rate of SLE in Caucasians is approximately 2 -8 cases per 100,000 individuals per year with a prevalence of between 15 and 50 cases per 100,000 individuals. SLE predominately affects women (female: male ratio = 5-9:1) and in particular ladies of childbearing age. One of the female sex-hormones, oestrogen, which has pro-inflammatory properties, is thought to be one of the elements capable for the female predominance[6] .It has been shown that some oestrogen containing oral contraceptives and pregnancy may cause the disease to flare and that disease activity may fluctuate with the menstrual cycle.

Table 4.1 - The American College of Rheumatology’s criteria for SLE, revised in 1982 and1997

Criterion	Definition
1. Malar rash	Fixed erythema over the malar eminence, tending to spare the nasolabial folds
2. Discoid rash	Erythematous raised patches with keratotic scaling and follicular plugging
3. Photosensitivity	Self-reported or observed rashes caused by unusual reaction to sunlight
4. Oral ulcers	Observed oral or nasopharyngeal ulceration

5.	Arthritis	Nonerosive arthritis involving two or more peripheral joints
6.	Serositis	Pleuritis or pericarditis
7.	Renal disorder	Persistent proteinuria >0.5g/day or cellular casts
8.	Neurologic disorder	Seizures or psychosis
9.	Hematologic disorder	Hemolytic anemia, leukopenia, lymphopenia or thrombocytopenia
10.	Immunologic disorder	Anti-DNA antibodies, anti-Sm antibodies, aPL antibodies, or false positive test for syphilis
11.	Antinuclear antibody	Abnormal titer of antinuclear antibody

4.1.2 Pima Indian Diabetes Dataset

People with type 1 diabetes must take daily insulin injections to survive. This form of diabetes It mainly affects children or young adults, however can occur at any age. Type 2 diabetes (otherwise called grown-up beginning diabetes or non-insulin-subordinate[7] .However, it primarily affects kids and youthful grown-ups doesn't produce enough insulin and/or doesn't use insulin properly (insulin resistance). This kind of diabetes for the most part happens in individuals over the age of 40 who are overweight and have a family background of diabetes, but it is now increasingly occurring in younger people, especially teenagers[8]. Type 2 diabetes (non-insulin-dependent) is diabetes is the most remarkable sort (90% to 95%), mainly occurring in adults but now also affecting children and adolescents. Type I (insulin-dependent) diabetes mainly affects children and young adults and is the least common form of diabetes (5% to 10%). Major risk factors for diabetes include obesity, high cholesterol, High blood pressure, no real work. As people age, the gamble of creating diabetes like wise increments [9]. Individuals who foster Gestational diabetes (called gestational diabetes) are bound to create full-blown diabetes sometime down the road. Poorly controlled diabetes can lead to many long-term complications, including heart disease, stroke, blindness, kidney failure, and vein infections..

The Pima Indian Diabetes Dataset is taken from the UCI AI Storehouse. The dataset has 768 instances with two types of questions testing whether a patient is positive or negative for diabetes[10].This dataset's patients were Pima Indian women living near Phoenix, Arizona, USA. This data set consists of 9 attributes as shown in Table 4.2.

Table 4.2 - Pima Indian Dataset

No.	Attribute	Description	Missing Values
1	pregnant	Number of times pregnant	110
2	glucose	Plasma glucose concentration (glucose tolerance test)	5
3	pressure	Diastolic blood pressure (mm Hg)	35
4	triceps	Triceps skin fold thickness (mm)	227
5	insulin	2-Hour serum insulin (mu U/ml)	374
6	mass	Body mass index (weight in kg/(height in m)^2)	11
7	pedigree	Diabetes pedigree function	0
8	age	Age (years)	0
9	diabetes	Class variable (test for diabetes)	0

Class Distribution: Class esteem 1 is deciphered as "tried positive for diabetes"

Class Value: 0 - Number of instances - 500

Class Value: 1 - Number of instances – 268

4.1.3 Liver Disorder Dataset:

Several disease states can influence the liver. A portion of these infections incorporate Wilson's disease, hepatitis (irritation of the liver), liver malignant growth, and cirrhosis (persistent irritation that at last advances to organ disappointment)[11]. Alcohol alters the metabolism of the liver and may have pervasive harmful effects if taken long-term. Hemochromatosis can cause liver problems.

Symptoms may include:

S.NO	SYMPTOMS
1.	jaundice
2.	Tendency to bruise or bleed
3.	Ascites
4.	Impaired brain function
5.	General failing health

4.2 Data Cleaning

First analyze whether there are missing qualities in the fields obtained in data cleaning [12]. If not addressed, missing values can alter the course what's more, course of the outcomes. Therefore, the first step is to handle missing qualities. Figure 4.1 Output Showing Missing Values

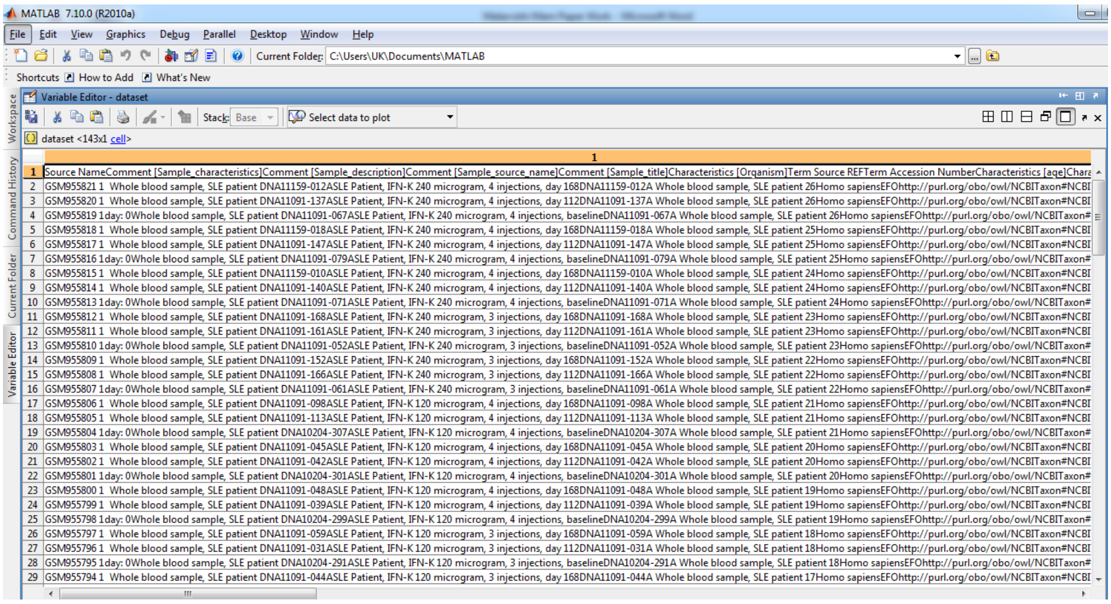

Fig. 4.1 - Missing Values Output

The avoidance of redundancy is followed by normalization of missing values. Fill the average mode value to filter rows. If 1 or 2 attribute values are missing, the median mode value is filled from the corresponding row value [13]. Once the attribute value is cleaned up, the attribute is reduced. As a final preprocessing step, irrelevant analysis attributes were removed. Attributes like Patient ID, URL, Array Index, Sample, etc. are not selected [14] . The data for the disease dataset obtained from CLD is selective and includes age, gender, test samples, disease activity, symptoms, severity, organs involved, tests performed and follow-up. Table 4.3 lists the CLD attributes used for the study and cleaned by preprocessing.

Table 4.3 – Cumulative Attributes

Attributes
Age
Gender
Test Sample
Disease Activity
Symptoms

Severity
Involved Organ
Tests Taken
Follow up

4.3 Decision Making

Decision tree (DT) is a non-parametric regulated learning technique for grouping and relapse. Decision trees create classification or regression models as a tree structure. Decision trees learn from data, using a set of if-then-else decision rules to approximate sinusoids. The further the tree, the more intricate the choice standards and the stricter the model[15]. Choice trees are quite possibly of the most mind-blowing known decision-making techniques, probably because of their inherent ease of outwardly conveying a decision or on the other hand set of decisions and their related vulnerabilities and results. Typically, this is the tree that produces the lowest cross-validation error. Decision trees are based on key factors of entropy and data gain In entropy, a choice tree is constructed hierarchical from the root hub and includes dividing the information into subsets containing occasions with comparable (homogeneous) values[16].If the sample is perfectly uniform, the entropy is zero; on the off chance that the example is equitably separated, the entropy is 1. Information gain entropy and information gain In entropy in entropy after splitting a dataset on one attribute. Building a choice tree is tied in with finding the trait that profits the best data gain (for example the most homogeneous branches).This paper uses the CART algorithm to form a decision tree according to the criteria listed in Table 4.4.

Table 4.4 –Decision Tree Criteria

Attribute	Condition	Decision Value
Age	<=30 >=31 and <=60 >61	0 1 2
Gender	Female Male	0 1
Test Sample	Blood Plasma Urine	0 1 2
Disease Activity	Mild Moderate Severe	0 1 2
Symptoms	None malar rash discoid rash photosensitivity oral ulcers non erosive arthritis pleuritis renal disorders neurologic disorder hematologic disorder immunologic disorders antinuclear antibody	0 1 2 3 4 5 6 7 8 9 10 11
Severity	Low Medium High	0 1 2
Involved Organ	none Skin Joints	0 1 2

	Musculoskeletal Blood Brain Lung Central Nervous System Vascular Eyes Heart Pulmunory Gastrointensional Mouth extremities	3 4 5 6 7 8 9 10 11 12 13 14
Tests Taken	None AntiNuclear Antibody Complete Blood Count Chest X-ray Kidney biopsy Urinalysis Rheumatoid test facts Liver function blood test Erythrocyte Sedimentation Rate	0 1 2 3 4 5 6 7 8
Follow up	Regular Occasional None	0 1 2

Therefore, the CLD dataset is separated into various properties. Process the entropy of each branch. Then add proportionally to get the total entropy of the division. Subtract the resulting entropy from the entropy before the division. The result is information gain, or a decrease in entropy [17] .The quality with the most noteworthy data gain is chosen as the choice hub, and the dataset is divided into relevant branches.

4.4 Ranking Results

In classification tasks, use ranking methods and develop new orders. This "grading-based approach" depends on the advancement positioning standards utilized in training. DMs are often asked to present preferences for a set of things as a positioning[18]. In this case, it's the severity range of symptoms. In the event that another standard is brought into a bunch of options, and the significance of that basis relies upon the quantity of other options and the strength of their positioning, the positioning of a bunch of choices might change. Thus, new alternatives can change the general request of the past set. This is achieved by pairwise correlation of options with criteria [19].It is especially useful in complex structures where the importance of criteria is not clear or where multiple criteria exist. The proposed work uses logical binary operations to sort sets according to rank. Table 4.5 lists the logical operations used in their classification**.**

Table 4.5 - logical operations for ranked output

Rule No.	Rule	Ranking
1	Age>1 && Gender>=0 && Test sample >1 && Disease Activity >=1 && Symptoms >1 && Severity >=1 && Involved Organ >2 && Test Taken>=1 && Follow up <=1	#1
2	Age>1 && Symptoms >1 && Severity >=1 && Involved Organ >2 && Test Taken>=1 && Follow up <=1	#2
3	Symptoms >1 && Severity >=1 && Involved Organ >2 && Test Taken>=1 && Follow up <=1	#3
4	Age>1 && Gender>=0 && Test sample >1 && Disease Activity >=1	#4

5	Age>1 && Gender>=0 && Test sample >1 && Disease Activity >=1 && Symptoms >1	#4
6	Age>1 && Symptoms >1 && Severity >=1 && Involved Organ >2 && Test Taken>=1	#5
7	Gender>=0 && Test sample >1 && Disease Activity >=1 && Symptoms >1 && Severity >=1 && Involved Organ >2	#5
8	Age>1 && Gender>=0 && Test sample >1 && Disease Activity >=1 && Symptoms >1 && Follow up <=1	#5
9	Test sample >1 && Disease Activity >=1 && Symptoms >1	#6
10	Symptoms >1 && Severity >=1 && Involved Organ >2	#6
11	Involved Organ >2 && Test Taken>=1 && Follow up <=1	#6
12	Age>1 && Test Taken>=1 && Follow up <=1	#6

The rules listed above are used to classify attributes. Here, the maximum severity of level 1 is achieved in rule 1, where the main parameters are age, gender, test sample; disease activity, symptoms, severity, organs affected, tests performed and follow-up are 1 or more [20]. Secondary parameters for Rule 2 include age, symptoms, severity, organs involved, tests performed, and follow-up. Rule 3 includes symptoms, severity, organs involved, tests performed and follow-up. Therefore, these 12 rules are applied sequentially in the CLD database, performing the logical operator AND for the severity attribute, performing the logical operator XOR for the normal attribute, and finally transforming the data into a range containing the severity.

4.5 Clustering

Bunching is the errand of partitioning a populace or data of interest into gatherings with the end goal that data of interest in a similar gathering are generally like different data of interest in a similar gathering and generally not the same as data of interest in different gatherings[21]. It is fundamentally an assortment of items in view of similitudes and contrasts between objects. Divide the results into three groups, grouping the values according to Table 4.6 lists the clustering results.

Table 4.6 –Clustering

Groups	Severity
Group 1	Severe
Group 2	Medium
Group 3	Mild

Clustering algorithms are powerful techniques for AI on unaided information. These algorithms are very powerful when applied to different machine learning problems and have been applied in different scenarios to help gain new insights into the problem [21]. Classification algorithms can use grouped output when classifying multivariate and complex data, such as symptom data. The proposed cluster grouping is shown in Figure 4.2.

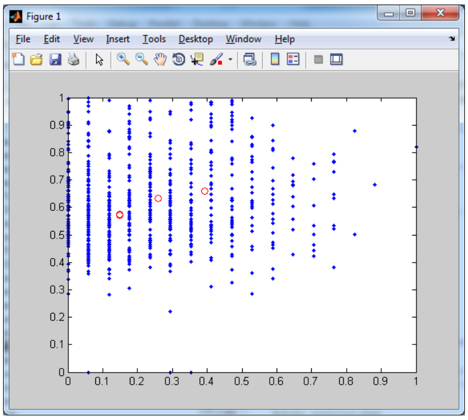

Fig. 4.2– Centroid Grouping

V. EXPERIMENTAL RESULTS

Improved Healthcare Prediction using Artificial Intelligence Approach (IHPAI) is an important algorithm type for predictive modeling machine learning[22]. The goal of creating a model is to anticipate the worth of a target (or dependent variable) based on the values of multiple inputs (or independent variables). IHPAI is a machine learning method for predicting disease and classifying patients with active disease from SLE datasets. Purify and select data for classification to improve accuracy [23]. IHPAI uses decision trees for decision making and implements CART to its advantage. In IHPAI, the objective variable is clear cut and trees are utilized to identify the "classes" the objective variable may belong to. In regression trees, the objective variable is persistent and the tree is utilized to anticipate its worth.[24]. The IHPAI algorithm consists of a series of questions, the answers to which determine what the next question (if any) is. The consequence of these inquiries is a tree structure with terminal hubs toward the end, so, all in all there are no more inquiries. The primary components of IHPAI are

• Rules for splitting data at nodes according to variable values;

• Stopping rules to decide when a branch terminates and cannot be further divided; and

• Prediction of the target variable for each terminal node.

Algorithm: IHPAI

Input: The number of clusters and cluster sub set, CLD dataset with n entities.

Output: A set of optimal clusters K_i.

Step 1: Initialize dataset $\sum$ (F) = { f_1,f_2,f_3 … f_n} attributes.

Step 2: Identify the Outliers in the considered column $\sum$ (F’) = { f_1’,f_2’,f_3’ … f_n’}

Step 3: Repeat, formulate the rules for identifying the similar attributes.

Step 4: do until, Identify the frequent itemsets.

Step 5: Specify the threshold Mean, proportion value.

Step 6: Identify the K initial mean vector from the attributes

Step 7: Identify the distance between f_i attributes and the centroid value f_j.

Step 8: Recalculate until new centroid f_j identified.

Step 9: Identify the end convergence.

Step 10: Fin the neighborhood active attribute rule set.

Step 11: Generate recommendations from most frequent itemset.

Step 12: Identify the disease threshold mean prediction value.

Table 5.1 - comparative performances of IHPAI by ml techniques

Algorithms	Performance Metrics	All feature set	After Cluster set
CART	Sensitivity	97.33	98.1
	Specificity	97.66	98.7
	Accuracy	97.19	98.1
	Elapsed Time (Sec)	7.451	6.903
K-Means	Sensitivity	93.2	95.33
	Specificity	93.1	96.7

	Accuracy	93.5	95.89
	Elapsed Time (Sec)	9.826	8.910
Decision Tree	Sensitivity	96.1	97.9
	Specificity	97.3	98.12
	Accuracy	96.88	97.92
	Elapsed Time (Sec)	8.663	7.933
Back Propagation	Sensitivity	97.13	98.72
	Specificity	97.86	98.11
	Accuracy	97.45	98.45
	Elapsed Time (Sec)	6.226	5.439
IHPAI	Sensitivity	98.33	98.88
	Specificity	98.66	99.17
	Accuracy	98.45	99.16
	Elapsed Time (Sec)	5.955	5.308

To verify the presentation of the proposed IHPAI, a bottom-up analysis of the groups it forms is performed.[25]. From Table 5.1, it is obvious that IHPAI performs better than CART (97.19), K-Means (93.5) and decision tree (96.8), ANN back propagation (97.45) in the expectation precision of CLD sickness, obtaining 98.45 points dataset.

VI. CONCLUSION

This work proposes three techniques capable of identifying health diseases from repository data. However, limitations of presently accessible tests have incited the advancement of different techniques that might offer more prominent explicitness, prognostic worth, cost-effectiveness, and other advantages. This research work proposes and demonstrates that health data can be predicted from patient datasets. The work proposes and demonstrates the need for an innovative technology to detect chronic diseases based on test results. This research work demonstrates ways to overcome these limitations using decision trees and rule-based techniques, helping to work on the productivity and accuracy of predictions. Furthermore, the proposed technique tries to eliminate the pitfalls of automatic prediction of health diseases. In the future, this technique could be extended to extract predictive data for other diseases, as long as clinicians determine the criteria or parameters in the test results.

REFERENCES

[1] Kishor, Amit, and Chinmay Chakraborty. "Artificial intelligence and internet of things based healthcare 4.0 monitoring system." *Wireless personal communications* 127.2 (2022): 1615-1631.

[2] Krishnamoorthy, Sreelakshmi, Amit Dua, and Shashank Gupta. "Role of emerging technologies in future IoT-driven Healthcare 4.0 technologies: A survey, current challenges and future directions." *Journal of Ambient Intelligence and Humanized Computing* 14.1 (2023): 361-407.

[3] Dwivedi, Ruby, Divya Mehrotra, and Shaleen Chandra. "Potential of Internet of Medical Things (IoMT) applications in building a smart healthcare system: A systematic review." *Journal of oral biology and craniofacial research* 12.2 (2022): 302-318.

[4] Vimont, Alexandre, Henri Leleu, and Isabelle Durand-Zaleski. "Machine learning versus regression modelling in predicting individual healthcare costs from a representative sample of the nationwide claims database in France." *The European Journal of Health Economics* 23.2 (2022): 211-223.

[5] Rastogi, Rashi, and Mamta Bansal. "Diabetes prediction model using data mining techniques." *Measurement: Sensors* 25 (2023): 100605.

[6] Abdollahi, Jafar, and Babak Nouri-Moghaddam. "Hybrid stacked ensemble combined with genetic algorithms for diabetes prediction." *Iran Journal of Computer Science* 5.3 (2022): 205-220

[7] Ghaffar Nia, Nafiseh, Erkan Kaplanoglu, and Ahad Nasab. "Evaluation of artificial intelligence techniques in disease diagnosis and prediction." *Discover Artificial Intelligence* 3.1 (2023): 5.

[8] Kapoor, Bhaskar, Bharti Nagpal, and Meshal Alharbi. "Secured healthcare monitoring for remote patient using energy-efficient IoT sensors." *Computers and Electrical Engineering* 106 (2023): 108585.

[9] Sim, Sungho, and Myeongyun Cho. "Convergence model of AI and IoT for virus disease control system." *Personal and Ubiquitous Computing* 27.3 (2023): 1209-1219.

[10] Yang, Christopher C. "Explainable artificial intelligence for predictive modeling in healthcare." *Journal of healthcare informatics research* 6.2 (2022): 228-239.

[11] Guo, Chonghui, and Jingfeng Chen. "Big data analytics in healthcare." *Knowledge Technology and Systems: Toward Establishing Knowledge Systems Science*. Singapore: Springer Nature Singapore, 2023. 27-70.

[12] Alanazi, Abdullah. "Using machine learning for healthcare challenges and opportunities." *Informatics in Medicine Unlocked* 30 (2022): 100924.

[13] Ozcan, Mert, and Serhat Peker. "A classification and regression tree algorithm for heart disease modeling and prediction." *Healthcare Analytics* 3 (2023): 100130.

[14] Jha, Ritesh, Vandana Bhattacharjee, and Abhijit Mustafi. "Increasing the prediction accuracy for thyroid disease: a step towards better health for society." *Wireless Personal Communications* 122.2 (2022): 1921-1938.

[15] Singh, Rameshwer, and Rajeshwar Singh. "Applications of sentiment analysis and machine learning techniques in disease outbreak prediction–A review." *Materials Today: Proceedings* 81 (2023): 1006-1011.

[16] Radanliev, Petar, and David De Roure. "Advancing the cybersecurity of the healthcare system with self-optimising and self-adaptative artificial intelligence (part 2)." *Health and Technology* 12.5 (2022): 923-929.

[17] Jimma, Bahiru Legesse. "Artificial intelligence in healthcare: A bibliometric analysis." *Telematics and Informatics Reports* (2023): 100041.

[18] Alshamrani, Mazin. "IoT and artificial intelligence implementations for remote healthcare monitoring systems: A survey." *Journal of King Saud University-Computer and Information Sciences* 34.8 (2022): 4687-4701.

[19] Hennebelle, Alain, Huned Materwala, and Leila Ismail. "HealthEdge: a machine learning-based smart healthcare framework for prediction of type 2 diabetes in an integrated IoT, edge, and cloud computing system." *Procedia Computer Science* 220 (2023): 331-338.

[20] Talaat, Fatma M. "Effective prediction and resource allocation method (EPRAM) in fog computing environment for smart healthcare system." *Multimedia Tools and Applications* 81.6 (2022): 8235-8258.

[21] Baucas, Marc Jayson, Petros Spachos, and Konstantinos N. Plataniotis. "Federated learning and blockchain-enabled fog-IoT platform for wearables in predictive healthcare." *IEEE Transactions on Computational Social Systems* (2023).

[22] Rehman, Arshia, Saeeda Naz, and Imran Razzak. "Leveraging big data analytics in healthcare enhancement: trends, challenges and opportunities." *Multimedia Systems* 28.4 (2022): 1339-1371.

[23] Mohapatra, Subasish. "A stacking classifiers model for detecting heart irregularities and predicting cardiovascular disease." *Healthcare Analytics* 3 (2023): 100133.

[24] Vimala Ithayan, J., and C. Sundar. "A Secured Healthcare Management and Service Retrieval for Society Over Apache Spark Hadoop Environment." *IETE Journal of Research* 69.2 (2023): 684-703.

[25] Oh, Wonsuk, and Girish N. Nadkarni. "Federated Learning in Health care Using Structured Medical Data." *Advances in kidney disease and health* 30.1 (2023): 4-16.

[26] Senthil, R., B. Narayanan, and K. Velmurugan. "Develop the hybrid Adadelta Stochastic Gradient Classifier with optimized feature selection algorithm to predict the heart disease at earlier stage." *Measurement: Sensors* 25 (2023): 100602.

[27] Verma, Navneet, Sukhdip Singh, and Devendra Prasad. "A review on existing IoT architecture and communication protocols used in healthcare monitoring system." *Journal of The Institution of Engineers (India): Series B* 103.1 (2022): 245-257.

[28] Baker, Stephanie, and Wei Xiang. "Artificial Intelligence of Things for Smarter Healthcare: A Survey of Advancements, Challenges, and Opportunities." *IEEE Communications Surveys & Tutorials* (2023).

[29] Prabh Deep, Gaurav Dhiman, and Rohit Sharma. "Internet of things for sustaining a smart and secure healthcare system." *Sustainable computing: informatics and systems* 33 (2022): 100622.

[30] Shrivastava, Prashant Kumar, Mayank Sharma, and Avenash Kumar. "HCBiLSTM: A hybrid model for predicting heart disease using CNN and BiLSTM algorithms." *Measurement: Sensors* 25 (2023): 100657.