IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 6, June 2021

DOI 10.17148/IJARCCE.2021.10640

Smart Health

Jonnalagadda Murali[1],Moparthi Chandra Sekhar[2], Kandula Dileep[3], M.Ayyavaraiah[4]

Under the guidance of Assistant Professor **M.Ayyavaraiah**

Department of computer science and engineering, KKR & KSR Institute of Technology & Sciences, Guntur, Andhra Pradesh, India

ABSTRACT: Data mining (DM) is an instrument of pattern detection and retrieval of knowledge from a large quantity of data. Many robust early detection services and other health-related technologies have developed from clinical and diagnostic evidence in both the DM and healthcare sectors. Artificial Intelligence (AI) is commonly used in the research and health care sectors. Classification or predictive analytics is a key part of AI in machine learning (ML). Present analyses of new predictive models founded on ML methods demonstrate promise in the area of scientific research. Healthcare professionals need accurate predictions of the outcomes of various illnesses that patients suffer from. Present statistical models of healthcare remedies have been scientifically reviewed. The uncertainty between statistical methods and ML has now been clarified. The study of related research reveals that the prediction of existing forecasting models differs even if the same dataset is used. Predictive models are also essential, and new approaches need to be improved. Here we predict the disease by using the algorithms.

1. INTRODUCTION

Today's world poses three problems in the healthcare sector: a shortage of medical practitioners, aging, and increased healthcare expenditure. The older the person, the more likely that they would be aged, sick, and require long-term care and prescription services. It is good news that healthcare services will integrate ML and AI, and thereby becoming smart disease prediction, through the increasing rise in computer power efficiency. This will likewise aim at fostering sustainable development. The efficiency of care aims to optimize the financial and social value of the health sector efficiently, without compromising the interest of our customers and our ability to deliver coverage in the future. Our main theme is the user will enter the attributes. By using different types of algorithm the result will be predicted with good accuracy. Logistic regression, Neural network, Adaboost are algorithms that we has used for predicting the disease.

2. LITERARTURE SURVEY

It might have happened so many times that you or someone yours need doctors help immediately, but they are not available due to some reason. The Health Prediction system is an end user support and online consultation project. Here we propose a system that allows users to get instant guidance on their health issues through an intelligent health care system online. The system is fed with various symptoms and the disease/illness associated with those systems. The system allows user to share their symptoms and issues. It then processes user's symptoms to check for various illnesses that could be associated with it. Here we use some intelligent data mining techniques to guess the most accurate illness that could be associated with patient's symptoms. In doctor module when doctor login to the system doctor can view his patient details and the report of that patient. Doctor can view details about the patient search what patient searched for according to their prediction. Doctor can view his personal details. Admin can add new disease details by specifying the type and symptoms of the disease into the database. Based on the name of the disease and symptom the data mining algorithm works. Admin can view various disease and symptoms stored in database. This system will provide proper guidance when the user specifies the symptoms of his illness.

3. EXISITING SYSTEM

Sometimes the situation occurs when you need the doctor's help immediately, but they are not available due to some reason. This system allows the users to get analysis on the symptoms they give for predicting the disease they are suffering from. User will be asked to enter the symptoms, then system will processes those symptoms for various illness or disease that user could be a liked with. In this system we use some techniques of data mining to guess the most accurate diseases or illness that could be related with patient's symptoms. If the system is unable to provide solutions, it informs the user about probable disease they have. If the user symptoms does not exactly match with any disease in our database, shows the diseases user could probably have judging the input symptoms. This system tends to replace the existing system for going to the doctor for getting diagnosis on illness you are suffering from to a smart solution where you get instant diagnosis on entering symptoms in the system.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 6, June 2021

DOI 10.17148/IJARCCE.2021.10640

4. PROPOSED SYSTEM

To beat the downside of existing framework we have created smart health prediction System. A framework checks a patient at initial level and proposes the possible diseases. It begins with getting some information about manifestations to the patient. The user has to enter attributes to know about the disease. By using some algorithms in Data Mining it became very to find the disease.

BENEFITS OF PROPOSED SYSTEM

User will specify the symptoms caused due to his illness. System will ask certain question regarding his illness and system predict the disease based on the attributes specified by the patient. User will give feedback this will be reported to the admin. The Admin will get notification how many people had accessed the system and what all are the diseases predicted by the system.

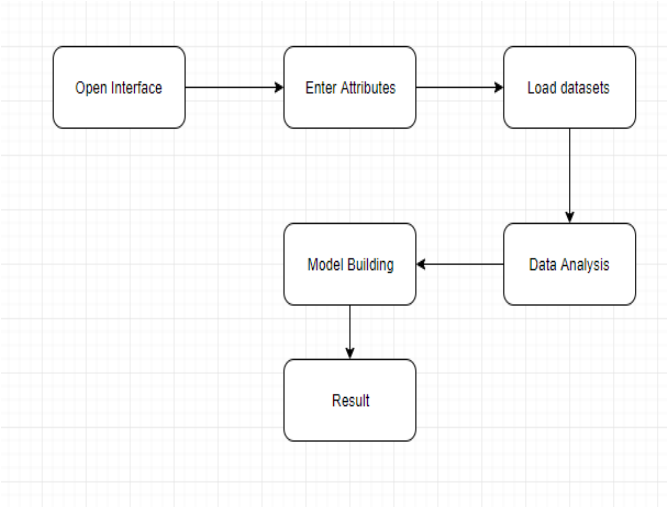
5. DESIGN

Design is a multi step process that focuses on data structure, Software architecture, procedural details and interface between modules. The design process also translates the requirements into the presentation of software that can be accessed for quality before coding begins.

Computer software design changes continuously as new methods; better analysis and broader understanding evolved. Software design at a relatively early stage in its revolution. Therefore, software design methodology lacks the depth, flexibility and quantitative nature that are normally associated with more classical engineering disciplines. However, the techniques for software design do exist, criteria for design qualities are available and design notation can be applied.

The purpose of the design phase is to plan a solution of the problem specified by the requirements document. The design of a system is perhaps the most critical factor affecting the quality of the software. It has a major impact on the project during later phases, particularly during testing and maintenance.

THE ARCHITECTURAL DESIGN:



6. MOTIVATION

Here we are used logistic regression , neural networks and Ada boost classifier for the prediction of diseases. We use this to predict the diseases . here in our system first we train the model and test the model and then if we give inputs it predicts the diseases . here we predict the autisum and diabetes diseases by using the above algorithms. The motivation behind this is we use the different algorithms to predict the diseases already we have previous prediction methods but that is not accurate. But in our system the accuracy increases. We predict the diseases with most accurately.

Logistic regression is a supervised learning classification algorithm used to predict the probability of a target variable. The nature of target or dependent variable is dichotomous, which means there would be only two possible classes.

In simple words, the dependent variable is binary in nature having data coded as either 1 (stands for success/yes) or 0 (stands for failure/no).Mathematically, a logistic regression model predicts P(Y=1) as a function of X. It is one of the simplest ML algorithms that can be used for various classification problems such as spam detection, Diabetes prediction, cancer detection etc.

AdaBoost algorithm, short for Adaptive Boosting, is a Boosting technique that is used as an Ensemble Method in Machine Learning. It is called Adaptive Boosting as the weights are re-assigned to each instance, with higher

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 6, June 2021

DOI 10.17148/IJARCCE.2021.10640

weights to incorrectly classified instances. Boosting is used to reduce bias as well as the variance for supervised learning. It works on the principle where learners are grown sequentially. Except for the first, each subsequent learner is grown from previously grown learners. In simple words, weak learners are converted into strong ones. Adaboost algorithm also works on the same principle as boosting, but there is a slight difference in working **NEURAL NETWORK** A neural network is a series of algorithms that endeavors to recognize underlying relationships in a set of data through a process that mimics the way the human brain operates. In this sense, neural networks refer to systems of neurons, either organic or artificial in nature. Neural networks can adapt to changing input; so the network generates the best possible result without needing to redesign the output criteria. The concept of neural networks, which has its roots in artificial intelligence, is swiftly gaining popularity in the development of trading syatems.

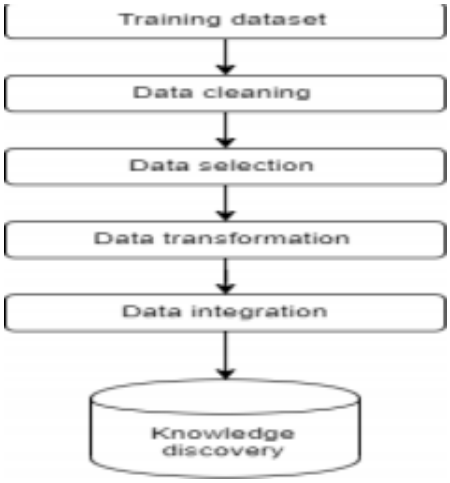
The two diseases we are predicting are:
A.Autisum
B.Diabetes

A. Autisum::

Here first we will collect the datasets from the kaggile website and save in the word document. Now we will train the datasets by giving the input as csv file format and train the datasets and now we test the datasets whether they are correct or not. After testing we give input and test the this is prediction. Here we use neural network for prediction of this disease.

B. Diabetes

Diabetes is the mostly occurred disease now a days this is not having any medicine to cure and get rid out of it throught out the life we are going to use the medicine so this is very dangeours disease now a days the world is facing. So if we find this earlier it will be low and it don't effect more so to predict this disease we trained a model and it predicts 74% accuracy which is high because the model cannot predict with 100% accuracy. Here we are going to use the logistic regression and ada boost classifier to predict this disease. First we collect the datasets and train them and test them. In the datsets we separate the datasets for training and testing. After train the model we test whether it is predicting correct or not.



This is the process for the prediction of diseases.
Here first we collect the datasets and then train the datasets and data cleaning removing the unwanted data and data selection and now the data transformation we are converting to our understandable data and data integration and knowledge discovery is final.

Algorithmic design

Step-1: Start
Step-2: Open Interface and Load datasets
Step-3: Identify and remove the missing data in observation function
Step-4: Exploratory data analysis (attributes like Age, Bp, Glucose....etc)
Step-5: Model Building

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 6, June 2021

DOI 10.17148/IJARCCE.2021.10640

Step-6: Model Training and Evaluation
Step-7: Display result
Step 8: Stop.

7. CODING OUTPUT SCREENS

The coding screen output for the autisum disease :

```
Prediction Reults for Neural Network
0.9152542372881356
              precision    recall  f1-score   support

           0       0.97      0.89      0.93        37
           1       0.84      0.95      0.89        22

    accuracy                           0.92        59
   macro avg       0.91      0.92      0.91        59
weighted avg       0.92      0.92      0.92        59

A1_Score             0
A2_Score             1
A3_Score             1
A4_Score             1
A5_Score             0
                    ..
relation_?           0
relation_Parent      1
relation_Relative    0
relation_Self        0
relation_self        0
Name: 79, Length: 96, dtype: int64
[0]
```

Here we got the output the accuracy of 91 to 92% accuracy for the autisum disease which is used for prediction of the human behavior.

The coding screen output for diabetes disease ;

This is for logistic regression algorithm.

```
--------- LogisticRegression() -----------
----TRAINING REPORT------
              precision    recall  f1-score   support

           0     0.7641    0.8886    0.8217       350
           1     0.7000    0.4866    0.5741       187

    accuracy                         0.7486       537
   macro avg     0.7321    0.6876    0.6979       537
weighted avg     0.7418    0.7486    0.7355       537

AUC-ROC Score:  0.77

----VALIDATION REPORT------
              precision    recall  f1-score   support

           0     0.8232    0.9000    0.8599       150
           1     0.7761    0.6420    0.7027        81

    accuracy                         0.8095       231
   macro avg     0.7996    0.7710    0.7813       231
weighted avg     0.8067    0.8095    0.8048       231

AUC-ROC Score:  0.77
```

This is for adaboost classifier algorithm:

```
--------- AdaBoostClassifier() ----------
----TRAINING REPORT------
              precision    recall  f1-score   support

           0     0.8643    0.8914    0.8776       350
           1     0.7841    0.7380    0.7603       187

    accuracy                         0.8380       537
   macro avg     0.8242    0.8147    0.8190       537
weighted avg     0.8363    0.8380    0.8368       537

AUC-ROC Score:  0.74

----VALIDATION REPORT------
              precision    recall  f1-score   support

           0     0.8176    0.8067    0.8121       150
           1     0.6586    0.6667    0.6585        81

    accuracy                         0.7576       231
   macro avg     0.7341    0.7367    0.7353       231
weighted avg     0.7590    0.7576    0.7582       231

AUC-ROC Score:  0.74
```

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 6, June 2021

DOI 10.17148/IJARCCE.2021.10640

8. CONCLUSION:

The data mining can play a vital role in disease prediction to design a smart health prediction system. In medical diagnosis, data mining has been widely used for predicting diseases through diagnosis. However, no single data mining algorithm is best suited to resolve the prediction issues for healthcare data sets. In conclusion, the combination of several data mining or hybrid version of data mining algorithm may be a better approach in designing health prediction system. The future research may be directed towards designing a better data mining based model that can address the healthcare with real-time healthcare datasets.

This study does not encompass the complete analysis of all existing data mining algorithms and real-time healthcare dataset. Besides, the proposed health prediction system is not built through the comparison of all the data mining algorithms available in literature. However, the future research may be directed towards the selection of the best suitable data mining algorithm through the analysis of all existing algorithms.

9. REFERENCES:

S. Chowdhury, "A Review Paper on Health Data Integration with Secured Record Linkage in National Health Data Warehouse and Privacy & Compatible Solution for its Privacy and Security Problems in Bangladesh and other developing countries," International Islamic University Chittagong (IIUC), 2017. [Online]. Available at: https://works.bepress.com/sayem-binsarwarchowdhury/2/download/. [Accessed 18th January 2021],

J. Tulsky, M. Beach, P. Butow, S. Hickman, J. Mack, R. Morrison, R. Street, D. White and K. Pollak, "A research agenda for communication between health care professionals and patients living with serious illness," JAMA internal medicine, vol. 177, no. 9, pp.1361-1366, 2017. [Online]. Available at: https://sci-hub.tw/https://jamanetwork.com/journals/jamainternalmedicine/article-abstract/2635330#. [Accessed 19th November 2020],

C. Kuziemsky, "Decision-making in healthcare as a complex adaptive system. Healthcare management forum," The Canadian College of Health Leaders, vol. 29, no. 1, pp. 4-7, 2016. [Online]. Available at: https://sci-hub.tw/https://doi.org/10.1177/0840470415614842#. [Accessed 20th November 2020],

K. Lakshmi, D. Ahmed and G. Kumar, "A smart clinical decision support system to predict diabetes disease using classification techniques," IJSRSET, vol. 4, no. 1, pp.1520-1522, 2018. [Online]. Available at: https://s3.amazonaws.com/academia.edu.documents/57021280/3909.pdf?response-content-disposition=inline%3B%20filename%3DA_Smart_Clinical_Decision_Support_System.pdf&X-Amz-Algorithm=AWS4-HMAC-SHA256&X-Amz-Credential=AKIAIWOWYYGZ2Y53UL3A%2F20191203%2Fus-east-1%2Fs3%2Faws4_request&X-Amz-Date=20191203T074834Z&X-Amz-Expires=3600&X-Amz-SignedHeaders=host&X-Amz-Signature=6ab09f94ad28907675bca9865b785e8c649eecccad6647d5696fd2233fde7859. [Accessed 20th November 2020],

M. Durairaj and V. Ranjani, "Data mining applications in healthcare sector: a study.," International Journal of Scientific & Technology Research, vol. 2, no. 10, pp. 29-35, 2013. [Online]. Available: http://www.ijstr.org/final-print/oct2013/Data-Mining-Applications-In-Healthcare-Sector-A-Study.pdf.

D. KS and A. Kamath, "Survey on Techniques of Data Mining and its Applications," International Journal of Emerging Research in Management & Technology, vol. 6, no. 2, 2017. [Online]. Available: https://pdfs.semanticscholar.org/b738/3df4705133a132f58104b514b80555fe78cb.pdf. [Accessed 8 February 2020].

V. Kirubha and S. Priya, "Survey on data mining algorithms in disease prediction," International Journal of Computer Trends and Technology (IJCTT), vol. 38, no. 3 , pp. 24-128, 2016. [Online]. Available: https://pdfs.semanticscholar.org/0d05/05b5b057333888de85043ae9b41259f07d85.pdf. [Accessed 8 February 2020].

B. Thakkar, M. Hasan and M. Desai, "Health care decision support system for swine flu prediction using naïve Bayes classifier," IEEE International Conference on Advances in Recent Technologies in Communication and Computing, pp. 101-105, 2010. [Online]. Available: https://sci-hub.tw/https://ieeexplore.ieee.org/abstract/document/5656881.

S. Palaniappan and R. Awang, "Intelligent heart disease prediction system using data mining techniques," IEEE/ACS international conference on computer systems and applications, pp. 108-115, 2008. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/4493524.

V. Krishnaiah, D. Narsimha and C. D.N.S., "Diagnosis of lung cancer prediction system using data mining classification techniques," International Journal of Computer Science and Information Technologies, vol. 4, no.1, pp. 39-45, 2013. [Online]. Available: https://pdfs.semanticscholar.org/22d9/ee58f155d7b706852105804faa72626e8ab1.pdf.