iJRASET
International Journal For Research in Applied Science and Engineering Technology

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 9 **Issue:** VII **Month of publication:** July 2021

**DOI:** https://doi.org/10.22214/ijraset.2021.36234

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429*
*Volume 9 Issue VII July 2021- Available at www.ijraset.com*

# Disease Prediction and Doctor Recommendation System using Machine Learning Approaches

Anand Kumar[1], Ganesh Kumar Sharma[2], U.M. Prakash[3]
[1,2] *UG Scholar,* [3]*Assistant Professor, School of Computing, SRM Institute of Science and Technology, Tamil Nadu, India*

***Abstract*— In our everyday life we go over numerous individuals who are experiencing some sort of Diseases. Prediction of disease is an integral part of treatment. In this project the disease is accurately predicted by looking at the symptoms of the patient where the patient can input his/her symptom and the system will predict the disease patient is suffering from. Classification Algorithms like the Naïve Bayes (NB), Random Forest, Logistic Regression and KNN have been broadly utilized to anticipate the Disease, where different accuracies were obtained. In corresponding to a particular Disease, for example, Heart Disease, Diabetes and so on is additionally anticipated by demonstrating "True" or "False" i.e. if an individual is having or not having that Disease. Prediction of such a system can have a very large potential in the medical treatment of the future. Once the Disease is predicted by the system, It then recommends which type of doctor to consult. In this paper, an audit of some new works identified with utilization of Machine Learning in expectation of disease is predicted. An interactive interface is built as front-end to facilitate interaction with the symptoms. The whole model is implemented using Django and is connected to the Django Server.***

***Keywords*— Disease Prediction, Machine Learning Classification, Django, Doctor Recommendation, Web Development, Python***

## I. INTRODUCTION

Machine learning is an arising approach that helps in determination of diseases using model information or past data. The Machine learning algorithm has two passes: Training and Testing. A Forecast of a disease by using a patient's symptoms and history machine learning technology is battling for past many years. Machine Learning technology gives a decent stage in the clinical field, so a medical care issues can be solved proficiently. We are applying machine learning to maintain complete hospital data Machine learning technology which allows building models to get rapidly examine information and convey results quicker. with the use of machine learning technology doctors can make good decision for patient diagnoses and treatment options, which prompts improvement of patient medical care administrations. Machine Learning technology gives a decent platform in medical field, so that a healthcare issues can be tackled proficiently.

The fundamental target of our project is to make the user Interface as a part of front end and connect it with ml models using Django python web framework. We will predict the disease in 2 variants i.e. in general and for specific disease. There is a need to study and make a framework which will make it simple for an end user to anticipate the constant sicknesses without visiting a specialist or a doctor for the treatment and also predicts the type of doctors they should visit.

Machine Learning has great power to analyse and cope up with different diseases so that prediction is more accurate and it is cost effective in the treatment.

## II. RELATED WORK

We have reviewed five papers from different other sources to make this project. We studied the relationship between different algorithms and their performance in different scenarios for the prediction of disease.

[1] Aim of this paper is to improve the savvy treatment using Machine Learning technology to simplify the decision support system. It is a comprehensive paper on the diagnosis of heart disease by monitoring a person's heartbeat. The framework grants you to set the requirements of your pulse. Subsequent to setting these limits an individual can begin to screen the heartbeat and at whatever point a person's heartbeat outperforms a particular level he get an admonition of high pulse and the danger of coronary failure or the heart attack. Author Ahmed M. Alaa and Senthil Kumar Mohan have experimented with a combination of different factors and obtained 88.7% accuracy with a random hybrid forest.

[2] This Paper deals with classic supervised binary classification where it is given a number of attributes in the Dataset. The dataset includes Plasma glucose concentration Blood pressure (mm Hg), Body mass index Age (years) etc. A number of elements all with certain features is used to identify people affected by the disease. To tackle the problem, we should investigate the information, do any necessary changes, apply ML algorithm, train a model, check the exhibition of the prepared model and repeat with different algorithms until we locate the most exact outcome. The improvement of software or sites it is critical to recognize the framework necessities by appropriately gathering expected information to connect with provider and customer.

[3] This Paper clarifies that there is a need to study and make a framework which will make it simple for an end-user to anticipate the perpetual disease without visiting a doctor or specialist for diagnosis. It is useful and simple to identify the Various Diseases through looking at Symptoms of patient's utilizing different strategies of Machine Learning Models. This section of the paper results the accuracy using different algorithms such as Decision Tree (DT) with accuracy of 90.2%, Random Forest (RF) with accuracy of 95.28% and NB with accuracy of 88.08%.

[4] This Paper clarifies that innovation has been further developed in wellbeing industry to give answers for the patients by giving ideas of trained professionals and facilities where to concede and which expert ought to be counseled for the particular disease. The medical care industry gathers information from the patient's data set by applying information mining and Machine Learning Techniques is utilized to foresee the illness where disease is the main source for the human's death from the previous years. The set of data is analyzed by using random forest algorithm and KNN algorithm where it approaches the problem with a specified question to analyze and find the solution between two or more independent variables and dependent variables. In prior days in the hospitals they need to perform different tests and should sit tight for one day for all the connected lab reports where it prompts fruitless and proficiency, precision is less. We have attempted the application in a viable manner in diminishing dimensionality, disposing of unimportant information, expanding exactness in simpler way patients can discover the medical clinics with great quality consideration of specialists.

[5] This Paper provided a heart disease prognosis using supervised Learning algorithms. The algorithms used include SVM, KNN and Naïve Bayes . Also data set contain 3000 objects with 14 features. From the huge writing looked into, it was seen that the maximum research utilized the disease dataset, which contains just 303 objects with 14 features. Naive Bayes shows the best results as it takes less time and high Accuracy of 86.6 while Decision Tree gives accuracy of 78.69% and KNN with the accuracy of 77.85%.

## III. EXISTING SYSTEM

In the existing system only have the specific disease prediction framework which predicts the disease based on medical parameters and not on the symptoms. There models trained on the datasets are quite old and not easily accessible everywhere so in most of the areas only the tradition health care systems exist. doctor recommendation system is not available with the prediction system and only a limited number of user can access at a time as the User Interface is not effective. It does not have mobile web support.

### *A. Disadvantage of Existing System*

*1)* Challenges in Finding the best specialists for a specific Disease.

*2)* User Accessibility is not good enough

## IV. PROPOSED SYSTEM

In this exploration we have discovered the answer for the issues looking in existing framework where we have proposed the exactness, dependability and productivity by building up the highlights of different Diseases where we will discover most of the regular illnesses in individual wellbeing and we have introduced in one application with expectation of the Diseases by examining the symptoms gathered from the patient's record and taking the Reliable dataset from Different Hospitals.

This framework is utilized to anticipate the greater part of the disease. It acknowledges the structured and textual data from the user. Since only prediction is not feasible, It is important to build user interactive interface in which there will be various sections of user interactive form to enter the symptoms and four models based on different ML Algorithms such as Naïve Bayes, Random Forest Regression, Logistic regression and KNN is build and in Parallel we have implemented Doctor Recommendation which will tell user to consult a specific type of Doctor after Predicting the disease. we have made authentication system including data collection Portal for collecting the datasets for building ML Models.

### *A. Advantage of Proposed System*

*1)* The system could reduce medical errors and improve patient results.

*2)* Simple method of getting to the application for Disease Prediction and Doctor Recommendation System

*3)* Application has various alternatives to settle on choice effectively on Diseases.

*4)* General Disease Prediction based on symptom and specific Disease Prediction based on Medical parameters gives the output on the same interface.


Figure 1-Flow chart Ml Model Builder

## V. SYSTEM ARCHITECTURE

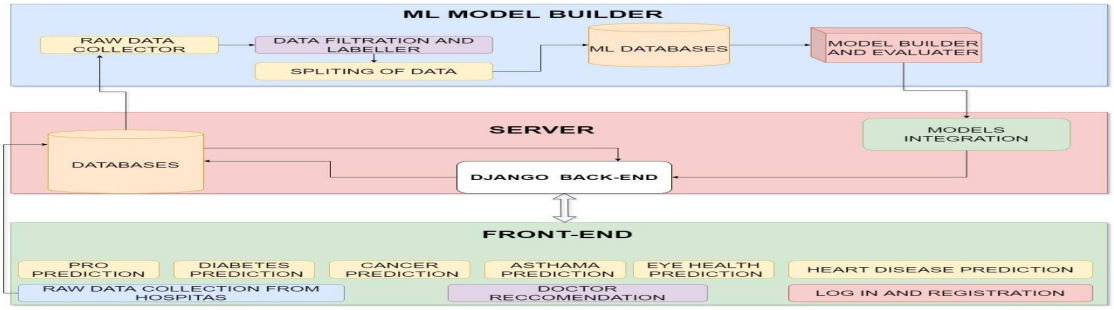



Figure 2: System Architecture

Architecture Diagram is mainly divided into three modules. First one is a ML model builder second one is Server and third one is Front end.

### *A. ML Model Builder*

There are various sections that is used for filtering of the data, that will be collected from the different hospitals for different diseases such as kidney disease, heart disease, cancer, hypertension, etc. and whatever the data will be collected, that will be the raw data and that data will be filtered and we have to label them. So for that process the ML model builder has been designed so that whatever the data has been collected in Central database is going to be recovered and then it will be processed into the data filtration and labeler where we will split the data for model building and splitting of data will be mostly 70(Training data)/30(Testing data) Or 80/20, based on the requirement and then it will be saved in ml database and after that it will passed to model builder and evaluator and that is the whole model building process and the testing is also there so we can say the model building part is totally consists of collection of raw data, filtration, feature Extraction, split the data set, fit the module analyze the data.

There are four different kind of models present in our project to predict the disease these are

Random Forest Classifier

Logistic Regression

Gaussian Naïve Bayes

KNN

And in evaluator we will test which model is performing better than other and whatever the result we will be getting that will integrate to the next module.

### *B. Server Module*

It mainly consists of Central database and Backend code written in the python Django framework. Here Models made in ML model builder will be integrated and all the functionality of the application interface for data collection from hospitals or disease prediction for patients will be written here. Here we will be collecting the raw data and storing the data from the different hospitals about different parameters of the required disease prediction. ML model integration will also be easy as we are using Python language for model building also. We will also be designing frontend for easy interface. So it will be more interactive.

### *C. Front End Module*

In front end module there are different tabs or we can say different sections. In every section there are different disease prediction module. For example, there is pro prediction tab where it is used for General prediction of health-wise problems, like

malaria, dengue, etc. and there are specific tabs also.

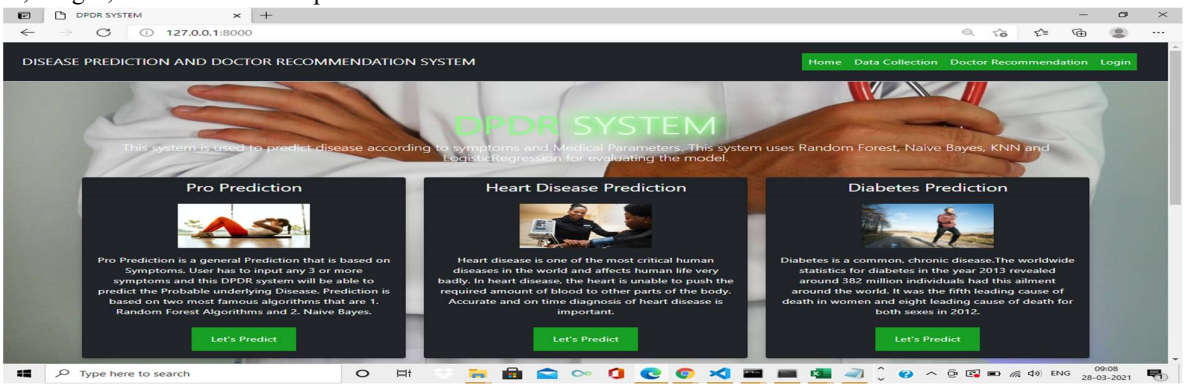



Figure 3- Homepage

1) *Homepage:* This is the Home page of Disease Prediction and doctor Recommendation System. Here we have brief information about Different Disease Prediction. Data Collection, Doctor Recommendation and Authentication system

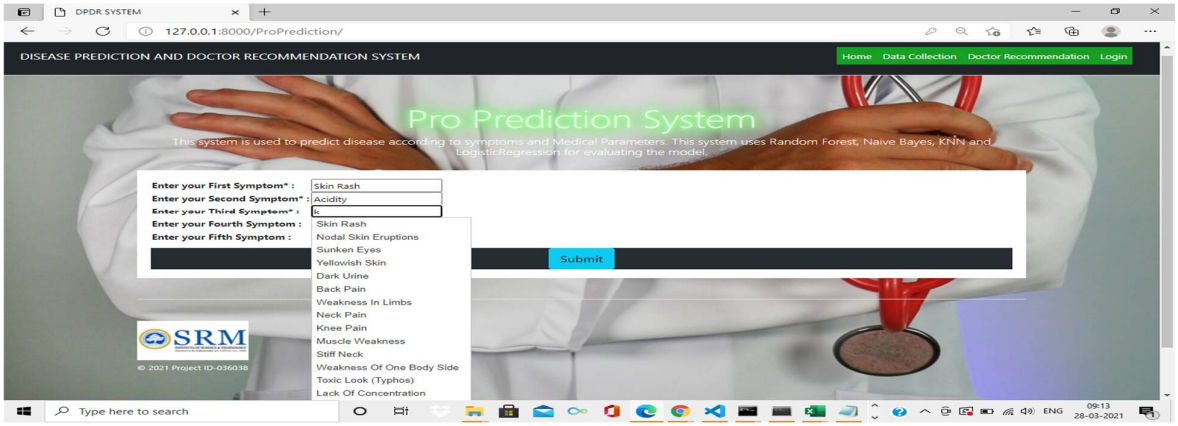



Figure 4- Pro Prediction Form

2) *Pro-Prediction Form:* Pro Prediction is a general Prediction that is based on Symptoms. User has to input any 3 or more symptoms and this DPDR system will be able to predict the Probable underlying Disease. Prediction is based on two most famous algorithms that are 1. Random Forest Algorithms and 2. Naive Bayes.

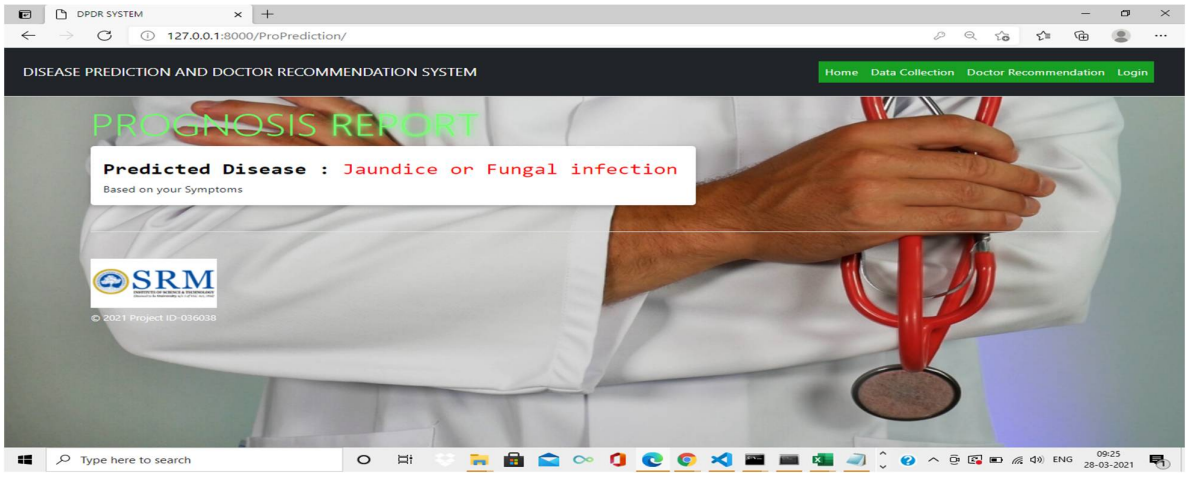



Figure 5- Pro Prediction output

3) *Pro-Prediction Output:* This gives the output of the disease which the user is suffering from based on the symptoms provided by the user. Once the user submit the symptoms the Disease is anticipated.

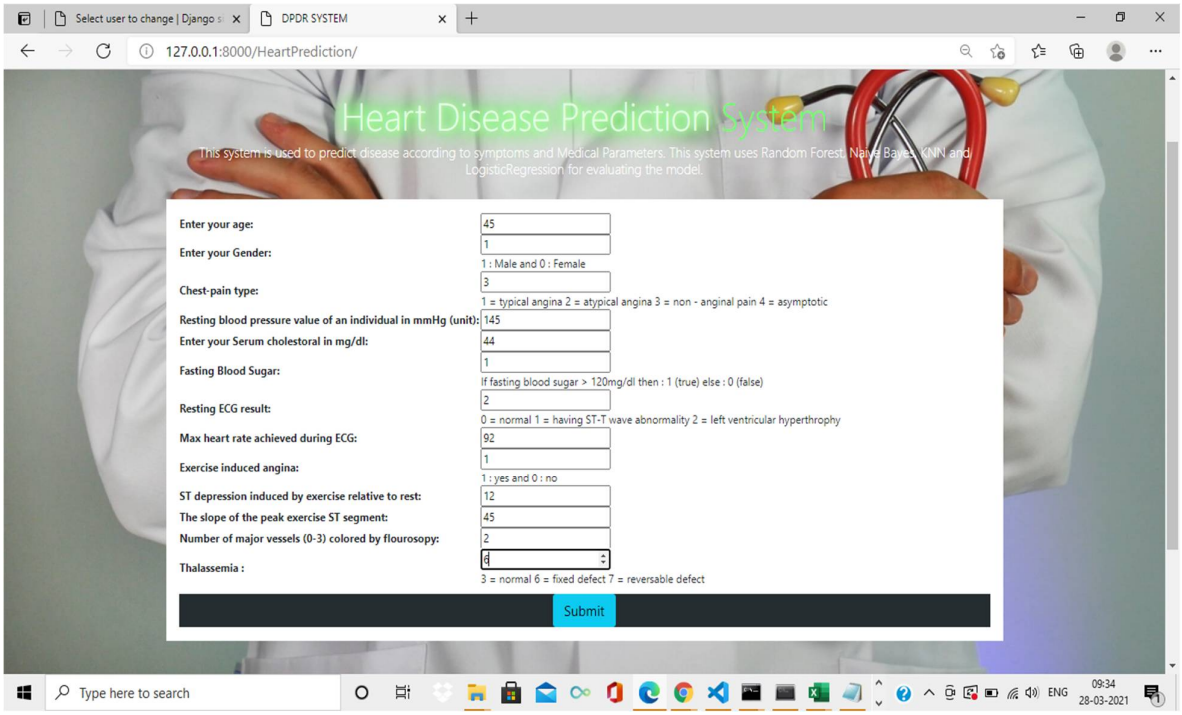



Figure 6- Heart Disease symptoms Input Form

4) *Heart Disease Symptoms Input Form:* This is user input form where the user Enters the symptoms and submits it. The system displays as “True” or “False” indicating that if the user is having or not having the Heart Disease.

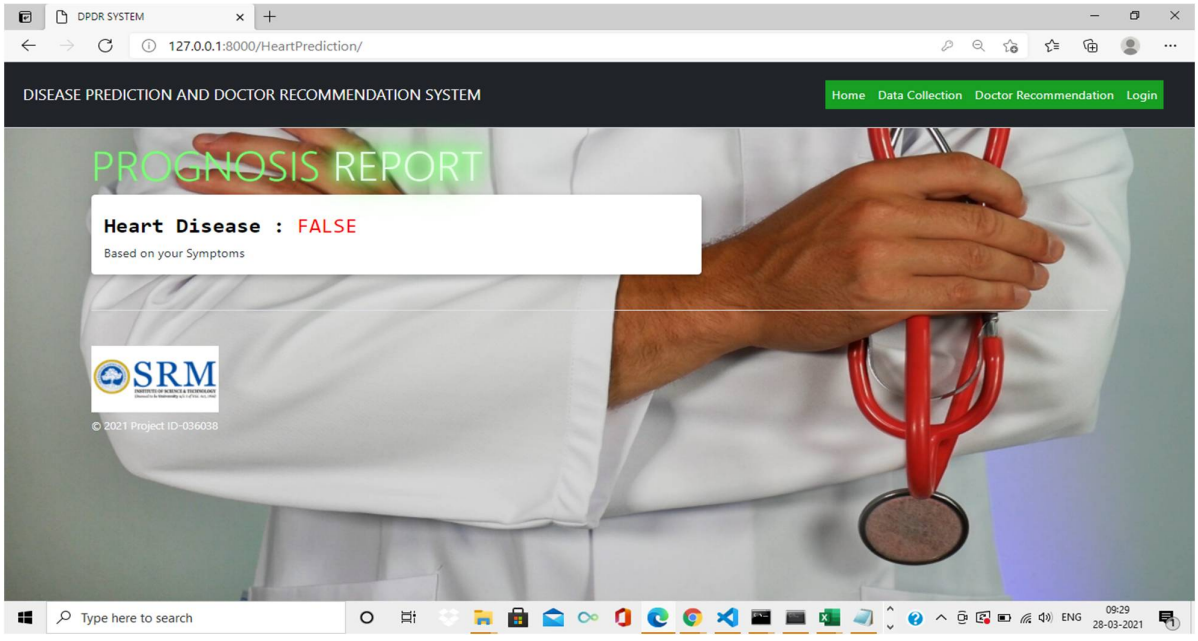



Fig 7- Heart Disease Prediction Output

5) *Heart Disease Prediction Results:* It predicts based on medical parameters which is provided by the user. Once the user submit the symptoms the Disease is anticipated as True or False.

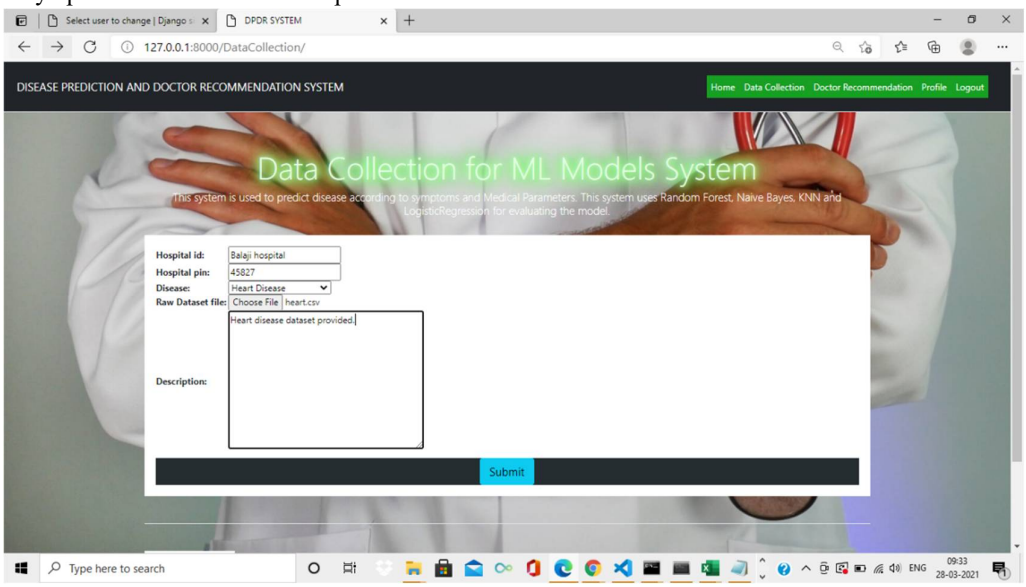



Fig 8- Dataset Collection Portal

6) *Data Collection Portal:* For different disease prediction, different data set is Required for that, there is data collection page. Raw Data about disease parameters is collected in forms through that page to the central database and then it is passed to the ML model builder.

7) *Doctor Recommendation:* There is a frontend part where we will have a tab for Doctor recommendation. So that if user wants to know the type of doctors they should visit based on the specific disease then this doctor recommendation system will help. This system will be based on disease prediction results. For example, if there is heart disease prediction than doctor recommendation system will recommend the cardiologist.

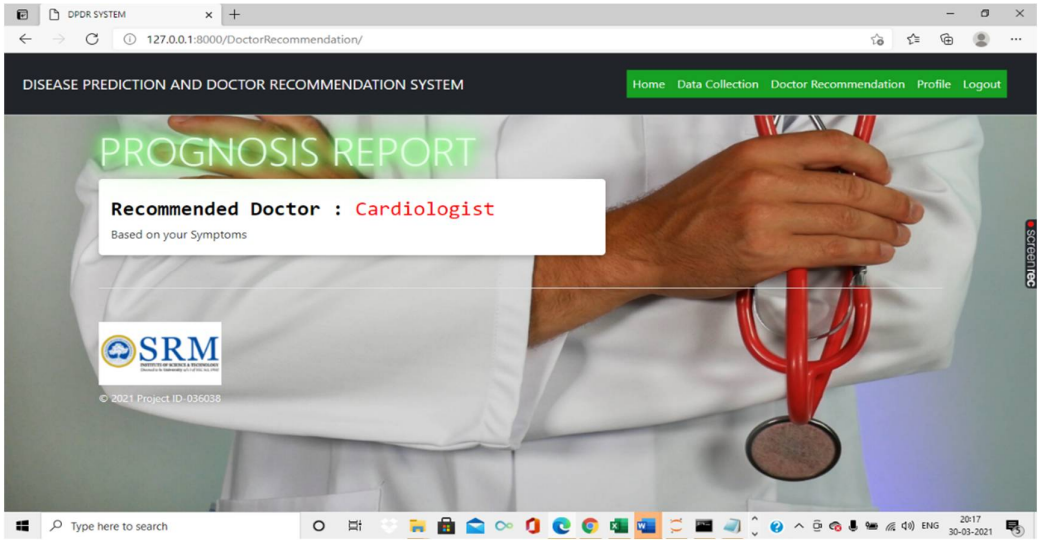



Fig 9- Doctor Recommendation System

## VI. DATASET AND MODEL DESCRIPTION

The dataset is used from various verified sources for different disease predictions. For heart disease prediction dataset has been

used from Cleveland Heart Disease dataset taken from the UCI repository. For diabetes prediction dataset has been used from Institute of Diabetes and Digestive and Kidney Diseases, and all the patients belongs to Pima Indian heritage, and are females of ages 21 and above. For Breast cancer prediction UCI Machine Learning Repository for breast cancer has been used.

### *A. Diabetes Attribute Information*

| | Pregnancies | Glucose | BloodPressure | SkinThickness | Insulin | BMI | DiabetesPedigreeFunction | Age | Outcome |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 6 | 148 | 72 | 35 | 0 | 33.6 | 0.627 | 50 | 1 |
| 1 | 1 | 85 | 66 | 29 | 0 | 26.6 | 0.351 | 31 | 0 |
| 2 | 8 | 183 | 64 | 0 | 0 | 23.3 | 0.672 | 32 | 1 |
| 3 | 1 | 89 | 66 | 23 | 94 | 28.1 | 0.167 | 21 | 0 |
| 4 | 0 | 137 | 40 | 35 | 168 | 43.1 | 2.288 | 33 | 1 |

Fig10- Diabetes Attributes

1) *Glucose:* Plasma glucose concentration over 2 hours in an oral glucose tolerance test
2) *SkinThickness:* Triceps skin fold thickness (mm)
3) *Insulin:* 2-Hour serum insulin (mu U/ml)
4) *Age:* Age (years)
5) *Pregnancies:* Number of times pregnant
6) *DiabetesPedigreeFunction:* Diabetes pedigree function
7) BloodPressure: Diastolic blood pressure (mm Hg)

### *B. Heart Disease Attribute Information*

```
df.head()
```

| | age | sex | cp | trestbps | chol | fbs | restecg | thalach | exang | oldpeak | slope | ca | thal | target |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 63 | 1 | 3 | 145 | 233 | 1 | 0 | 150 | 0 | 2.3 | 0 | 0 | 1 | 1 |
| 1 | 37 | 1 | 2 | 130 | 250 | 0 | 1 | 187 | 0 | 3.5 | 0 | 0 | 2 | 1 |
| 2 | 41 | 0 | 1 | 130 | 204 | 0 | 0 | 172 | 0 | 1.4 | 2 | 0 | 2 | 1 |
| 3 | 56 | 1 | 1 | 120 | 236 | 0 | 1 | 178 | 0 | 0.8 | 2 | 0 | 2 | 1 |
| 4 | 57 | 0 | 0 | 120 | 354 | 0 | 1 | 163 | 1 | 0.6 | 2 | 0 | 2 | 1 |

Fig 11- Heart Disease Attributes

| Name | Type | Description |
|---|---|---|
| Age | Continuous | Age in years |
| Sex | Discrete | 1 = male<br>0 = female |
| Cp | Discrete | Chest pain type:<br>1 = typical angina<br>2 = atypical angina<br>3 = non-anginal pain<br>4 =asymptomatic |
| Trestbps | Continuous | Resting blood pressure (in mm Hg) |
| Chol | Continuous | Serum cholesterol in mg/dl |
| Fbs | Discrete | Fasting blood sugar > 120 mg/dl:<br>1 = true<br>0 = false |
| Restecg | Discrete | Resting electrocardiographic results:<br>0 = normal<br>1 = having ST-T wave abnormality<br>2 =showing probable or define left ventricular hypertrophy by Estes'criteria |
| Thalach | Continuous | Maximum heart rate achieved |
| Exang | Discrete | Exercise induced angina:<br>1 = yes<br>0 = no |
| Old peak ST | Continuous | Depression induced by exercise relative to rest |
| Slope | Discrete | The slope of the peak exercise segment :<br>1 = up sloping<br>2 = flat<br>3= down sloping |
| Ca | Discrete | Number of major vessels colored by fluoroscopy that ranged between 0 and 3. |
| Thal | Discrete | 3 = normal<br>6= fixed defect<br>7= reversible defect |
| Diagnosis | Discrete | Diagnosis classes:<br>0 = healthy |

Fig 12- Heart Disease Attribute Description

*C. Pro Prediction Attribute Information*

```
df.head()
```

| | itching | skin_rash | nodal_skin_eruptions | continuous_sneezing | shivering | chills | joint_pain | stomach_pain | acidity | ulcers_on_tongue | ... | blackheads | scurri |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ... | 0 | |
| 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ... | 0 | |
| 2 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ... | 0 | |
| 3 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ... | 0 | |
| 4 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ... | 0 | |

5 rows × 133 columns

Fig 13- Pro Prediction Attributes

*D. Liver Prediction Attribute Information*

```
df.head()
```

| | Age | Gender | Total_Bilirubin | Direct_Bilirubin | Alkaline_Phosphotase | Alamine_Aminotransferase | Aspartate_Aminotransferase | Total_Protiens | Albumin | Album |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 65 | Female | 0.7 | 0.1 | 187 | 16 | 18 | 6.8 | 3.3 | |
| 1 | 62 | Male | 10.9 | 5.5 | 699 | 64 | 100 | 7.5 | 3.2 | |
| 2 | 62 | Male | 7.3 | 4.1 | 490 | 60 | 68 | 7.0 | 3.3 | |
| 3 | 58 | Male | 1.0 | 0.4 | 182 | 14 | 20 | 6.8 | 3.4 | |
| 4 | 72 | Male | 3.9 | 2.0 | 195 | 27 | 59 | 7.3 | 2.4 | |

Fig 14- Liver Disease Attributes

***Model Description***

*A. Logistic Regression*

*1)* It used to show the likelihood of a specific class or occasion existing like pass/fail, spam/not spam, win/lose, or solid/wiped out which is addressed by a pointer variable where the two entities are "0" and "1".

*2)* Logistic regression estimates relationship between the reliant variable and one or more independent variables by anticipating probabilities using a calculated Function.

*3)* It is the binary classification Algorithm where the values have category either 0 or 1. If new data Points needs to be fitted it should be in either of one category. If the data point is not between 0 and 1, we generally squish all data in the range [0,1] using Sigmoid Function.

*B. Random Forest Classification*

Random Forest Algorithm is the sort of Supervised learning calculation utilized for both the regression and Classification. This algorithm works on 4 basic steps –

*1)* It chooses random data samples from dataset.

*2)* It constructs decision trees for every sample dataset chosen.

*3)* At this step every predicted result will be compiled and voted on.

*4)* At last the most voted Prediction will be selected and be presented as result of classification.

*C. Naïve Bayes*

*1)* It is based on Bayes Theorem and used for classification problem.

*2)* It is a probabilistic classifier, which implies it predicts based on the likelihood of an object and can make fast Prediction. Probability for both the categories are checked for new data Point and the category with highest Probability gives the accuracy.

*D.* KNN

*1)* K - Nearest neighbour is a kind of directed learning calculation. It is an essential algorithm. It finds extensive use in pattern finding and data mining. It works by finding a pattern in data which links data to results and it improves upon the pattern recognition with every iteration.

*2)* It is Supervised Learning technique

*3)* KNN calculation at the preparation stage simply saves the dataset and after getting the new information then It classify the information that is Much like new Data.

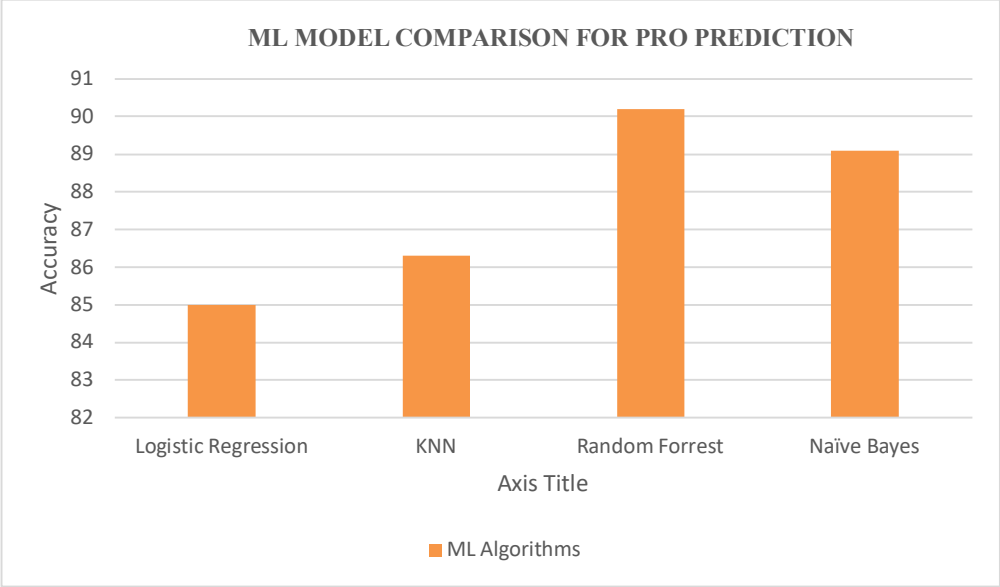
## VII. EVALUATION OF ML MODEL AND RESULTS

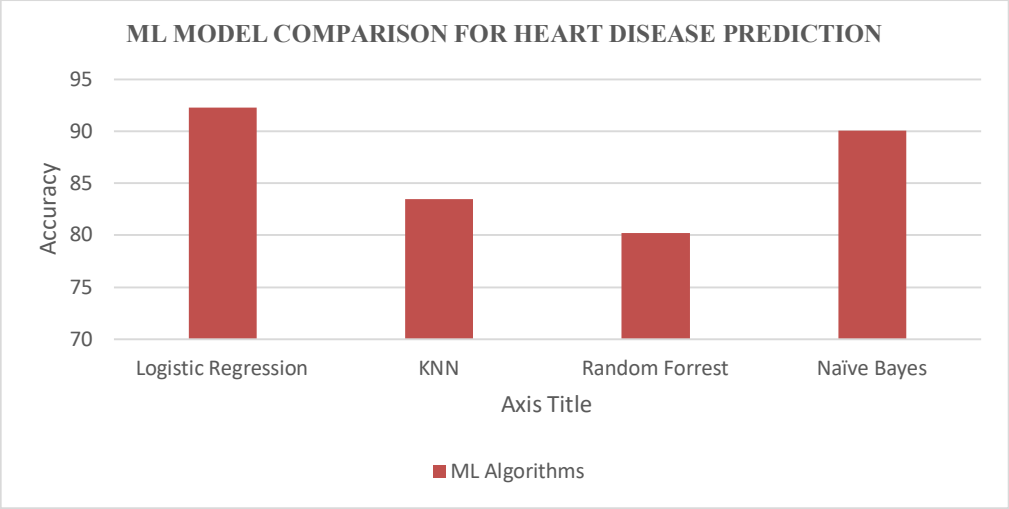
The Result for the Pro Prediction gives the output of the Disease from which the individual is suffering based on the symptoms provided by the individual. The algorithm taken for pro prediction is Random Forest Classifier and Naïve Bayes because of its highest accuracy i.e. 90.2%. In Parallel the anticipation for Specific Disease such as Heart Disease, Diabetes etc. is also predicted by indicating "True" or "False" i.e. having the disease or not having the Disease. The accuracy for the algorithm Logistic Regression is highest i.e. 92.3% followed by Naïve Bayes i.e. 90.1% in case of Heart Disease while the accuracy for the algorithm KNN is highest i.e. 74% followed by Naïve Bayes i.e. 72% in case of Liver Prediction Disease.

| | **Logistic Regression** | **Random Forest Classifier** | **KNN** | **Naïve Byes** |
|---|---|---|---|---|
| **Pro Prediction** | **85%** | **90.2%** | **86.3%** | **89.1%** |
| **Heart Disease** | **92.3%** | **80.2%** | **83.5%** | **90.1%** |
| **Diabetes** | **78%** | **70.6 %** | **71%** | **74%** |
| **Liver Disease** | **71.4%** | **71.4%** | **73.7%** | **73.1%** |

Table 1 – Models Accuracy

***Graphs For Model Comparison***






## VIII. CONCLUSION AND FUTURE SCOPE

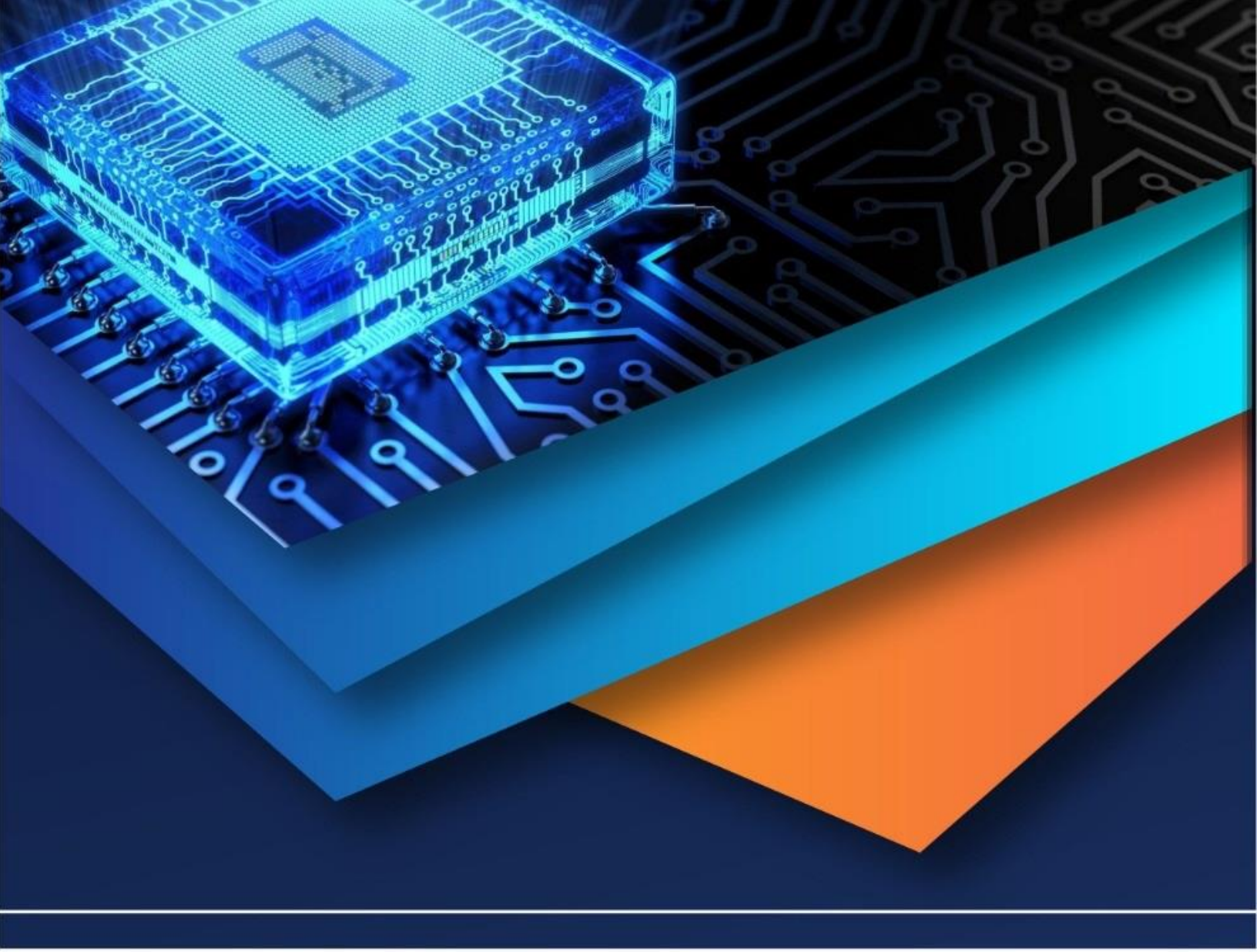
We set out to create a framework which can anticipate disease on the basis of symptoms given to it. Such a framework can reduce the surge at OPDs of clinics and lessen the responsibility on clinical staff. By creating this system we also added a way to store the data entered by the user in the database which can be used in future to help in creating better version of such system and while going through the past Papers we reviewed that this was never implemented before so this makes the Project more conclusive. Data analysis and machine learning Techniques have been utilized to gauge the disease events and have summed up to Determine each algorithm's output expectation yield and apply the appropriate method proposed for the area needed. It likewise has different visual portrayal of information gathered and results accomplished. It also has various visual representations of data collection and results achieved.

In the future we can also improve the model by implementing Deep Learning Algorithms by taking large and huge datasets directly from the Hospitals. We can Implement the whole Project inside the Android/Application so that it can be more accessible to the Users.

## REFERENCES

[1] Mrs. Shital. P. Chattar, Prathamesh Wadile , Bhushan Saraf , Omkar Kardile and Aditya Pawar. " IOT Based Heart Disease Prediction." International Journal for Research in Applied Science & Engineering Technology (IJRASET) (2021):2321-9653.
[2] Megha Vyas , Priya Kunwar, Vinod Joshi, Mihir Maheshwari and Mr. Girish Ameta. "Software for Disease Prediction." International Journal for Research in Applied Science & Engineering Technology(IJRASET) :2321-9653,2020.
[3] Yuvraj Nikhate and M. V. Jonnalagedda. "Survey On Heart Disease Prediction using Machine Learning." International Journal of Creative Research Thoughts(2020) :2320-2882.
[4] Raj H. Chauhan, Daksh N. Naik, Rinal A. Halpati, Sagarkumar J. Patel4 and Mr. A.D.Prajapati. " Disease Prediction Using Machine Learning."International Research Journal Of Engineering & Technology(IRJET):23950056,2020.
[5] Sandhya B , Monika K. J , Vyshnavi and Usman K. " Disease Prediction and Hospital Recommendation using Machine Learning Algorithms." International Journal for Research in Applied Science & Engineering Technology(IJRASET): 2321-9653,2020.
[6] Lamido Yahaya , Nathaniel David Oye and Etemi Joshua Garba. "A Comprehensive Review on Heart Disease Prediction Using Data Mining and Machine Learning Techniques. "American Journal of Artificial Intelligence:2639-9717 ,vol 4, 2020.
[7] Neelaveni, J., and MS Geetha Devasana. "Alzheimer disease prediction using machine learning algorithms." 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS). IEEE, 2020
[8] Terrada, Oumaima, et al. "Atherosclerosis disease prediction using Supervised Machine Learning Techniques." 2020 1st International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET). IEEE, 2020.
[9] Fitriyani, Norma Latif, et al. "HDPM: An Effective Heart Disease Prediction Model for a Clinical Decision Support System." IEEE Access 8 (2020): 133034-133050.
[10] Grampurohit, Sneha, and Chetan Sagarnal. "Disease Prediction using Machine Learning Algorithms." 2020 International Conference for Emerging Technology (INCET). IEEE, 2020.
[11] Verma, Anurag Kumar, Saurabh Pal, and Surjeet Kumar. "Prediction of skin disease using ensemble data mining techniques and feature selection method—a comparative study." Applied biochemistry and biotechnology 190.2 (2020): 341-359.
[12] Ramalingam, V. V., Ayantan Dandapath, and M. Karthik Raja. "Heart disease prediction using machine learning techniques: a survey." International Journal of Engineering & Technology 7, no. 2.8 (2018): 684-687.
[13] Uddin, Shahadat, et al. "Comparing different supervised machine learning algorithms for disease prediction." BMC medical informatics and decision making 19.1 (2019): 1-16.
[14] Ramalingam, V. V., Dandapath, A., & Raja, M. K. (2018). Heart disease prediction using machine learning techniques: a survey. International Journal of Engineering & Technology, 7(2.8), 684-687.
[15] Chen, Min, et al. "Disease prediction by machine learning over big data from healthcare communities." Ieee Access 5 (2017): 8869-8879.

[16] Vinitha, S., et al. "Disease prediction using machine learning over big data." Computer Science & Engineering: An International Journal (CSEIJ) 8.1 (2018).
[17] Sarker, Iqbal H. "Machine learning: Algorithms, real-world applications and research directions." SN Computer Science 2.3 (2021): 1-21.
[18] Karun, Ishita. "Comparative Analysis of Prediction Algorithms for Heart Diseases." Advances in Computer, Communication and Computational Sciences. Springer, Singapore, 2021. 583-591.
[19] Narayana, Ch V., et al. "Predicting Breast Cancer Using Machine Learning." Proceedings of International Conference on Computational Intelligence and Data Engineering. Springer, Singapore, 2021.
[20] Nkiruka, Odu, Rajesh Prasad, and Onime Clement. "Prediction of malaria incidence using climate variability and machine learning." Informatics in Medicine Unlocked 22 (2021): 100508.
[21] Praveen, A. Durga, et al. "Intelligent Liver Disease Prediction (ILDP) System Using Machine Learning Models." Intelligent Computing in Control and Communication. Springer, Singapore, 2021. 609-625.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY