iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** IV **Month of publication:** April 2023

DOI: https://doi.org/10.22214/ijraset.2023.50977

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: 7.538
Volume 11 Issue IV Apr 2023- Available at www.ijraset.com

Personal Healthcare Portal

Sanket Sawant[1], Kundan Badhe[2], Kunal Sanwadkar[3], Santosh Tamboli[4]
[1, 2, 3]*Students,* [4]*Asst. Prof., B.E. Information Technology, Vidyalankar Institute of Technology Mumbai, Maharashtra*

Abstract: ***Personal healthcare portal is a web-based application integrated with machine learning models which would be beneficial for symptom-based disease prediction and guiding about health tips. Apart from this it also can help users to communicate with medical experts. This project aims for enabling users to maintain their health with an application by testing and predicting various diseases to make personal awareness about remedial measures. Our application also focuses on saving time, time, efforts and money of users to clear their doubts about health issues with prior idea about health before/without appointing to doctors. Also, with some more specialization it can act as intermediate between different health centers and multiple users digitally. Most efficient Machine learning model amongst various supervised learning algorithms provide more accurate disease prediction based on symptoms and web portal make it easily accessible to information and service with relevant APIs. In general, it is an innovative approach to make services digital and online it.***
Keywords: symptom based- disease prediction, machine learning, web portal, APIs, supervised learning

I. INTRODUCTION

According to WHO, "*early diagnosis of diseases can minimize the risk of adverse effect of many diseases*". So, any technology which can provide preliminary estimation of any diseases according to person's current conditions may become a boon in healthcare industry. In this modern age due to fast and digital lifestyle, people are willing to find the solutions which are available online, in less time and from anywhere. This Project named 'PERSONAL HEALTHCARE PORTAL' which is a Machine Learning based Web application aims to provide that solution by developing fast and efficient algorithms and data-driven models for real-time processing of data to produce accurate prediction and results. This service can be a first, simple and quick step for healthcare before personally appointing with doctors. People from today's world are willing to examine their health status within just few clicks. We are building a web application that is committed to providing an effective, anonymous system which will guide people about their health by suggestions, predictions and interaction with experts. People will able to analyze and take decisions on basis of outcomes of provided system in smarter way by using their system anytime anywhere. Such applications can be a platform for information sharing, health-service management and online medical algorithms [1]. The main goal of this project is to analyze the data and accurate disease prediction. This application can create a local or global intermediaries between healthcare organizations, service providers and experts for exchange of medical solutions remotely. Implementation of decentralized application along with integration with multiple medical centers can create an online highly networked healthcare platform. Thus, we can infer that such system can be extended to the scope of medical service management, communication and information sharing platform about healthcare along with government as well as private medical sectors.

II. PROBLEM STATEMENT

People want to estimate possibility of causing disease based on current symptoms without any medical test. Also, they need a proper platform which is able to provide diet and workout suggestions. Apart from this easy contact with medical specialist on same place would be more suitable. So this need to build an User friendly application which will act as a quick responsive system for diagnosis of any short term or chronic disease based on basic input data from user. System would be able to provide remedial measures with tips and guidelines according to users need along with permitting user for contacting with appropriate authorities. Solution should be an effective means of platform which will provide virtual functionalities of health center by variety of ways.

III.LITERATURE REVIEW

A Through this section, we summarize some of the existing research work on ML based healthcare systems or web portals providing such solutions. In [2] Toqueer Ahmed, Saeed Mian Qaisar, proposed solution titled "Novel Web based multi class heart disease prediction using machine learning" at ICEEST on March 2023. Provided model is about web portal predicting heart disease with supervised ML algorithms along with artificial neural network. python Flask framework is used for web development and classifier used here are KNN classifier and simple vector classifier.

That has accuracy about 93.5%. Although it has better accuracy, system is only built for heart disease and due to deep -learning applications it consumes high time and power. Also large sized dataset is required.

As referred [3] on April 2020, Abhishek Singh and Dr. SPS Chauhan published their work regarding "Prediction and Analysis of Multiple diseases using Machine learning Techniques "from IRJET, India. They developed a system for prediction of various disorders with Machine learning algorithms like K nearest neighbor, decision tree and SVM having accuracy about 85%. Different algorithms are comparatively examined and selected. Still proposed model need to be efficient by increasing its accuracy. Separate prediction of each chronic disease on medical data inputs is implemented by training every disease dataset with each algorithm and only effective is finalized for estimation in presented project

In [4] on December 2019 Shahadat Uddin, Arif Khan, Mohammad Ali-Moni published a paper named "Comparing different supervised machine learning algorithms for disease predictions" by BMC informatics and decision making. Detailed description about comparison of multiple supervised ML models which are used in healthcare industry for various purpose are given in that research paper. Study about efficiency, use cases and statistical logic behind each model is summarized here. It concludes that SVM is most widely used whereas random forest is most accurate algorithm for disease prediction.

From [5] proposed work published by Harish Kumar Rajora, Narinder Singh Punn, Sonali Agarwal named "Web based disease prediction and recommender system" is web enabled system integrated with Database and machine learning algorithms like naïve bias, random forest and KNN. It is also capable of storing centrally data of users and personalization of recommendations along with common chronic disease prediction service.According to [7] published on April 2023, Nandikatla Chandrasekharan, Samineni Peddakrishna, proposed a paper called "Enhancing heart disease prediction accuracy through machine learning techniques and optimization", Mentioned paper is based on techniques for optimization of ML algorithms by increasing its efficiency for more correct disease predictions. Mainly it is based on some techniques like ensemble learning, n-fold cross validation, soft voting classifier which may increase accuracy up to 3-5 %

In [8] on August 2021, Arumugam K, Priyanka Shinde, Mohd Naved, Tatiana Gonzales published a work with title "multiple disease prediction using ML algorithms" to scientific committee of international conference on nanoelectronics, nanomaterials &nanotechnology. This work comprises of chronic disease estimation using Decision tree algorithm having accuracy of 90%. Datasets of various symptoms and comparative study of error rates of Naïve bias and SVM model with decision tree are remarkable points.

As [10] on March 2020, Amit Juyal, Chetan Pandey, Janmejay pant, Dr. Ankur Dhumka, published their reseach paper which is "performance analysis of supervised ML on medical dataset" that briefly describes about various performance measures of variety of supervised ML algorithms based on evaluation parameters like accuracy, precision, recall, ROC and F-1 score. This can help for evaluating and summarizing various models and selecting best suited amongst them according to their efficiency.

Apart from Technical publications some commercial online e-com portal like [15] "Myhealthportal" developed by Edward R Sykes, Syed Tanbeer for assisting homecare and clinic-based healthcare service like, online appointment scheduling, monitoring, and information sharing. And well known "Arogya setu" specially build by National Informatics Centre, GOI for online medical service management. But this are fully fledged, well-developed applications at high level with high cost. That's why it would be ideal for developing simple solution at minimum cost providing multiple services at single platform.

As mentioned [20] "Smartcare" developed by Stepheny Lucas, Mitali Desai, Amisha Khot, Sincee Harriet, Nilambari Narkar deals with implementation of healthcare system which is capable of disease prediction using ML technique and searching for nearby hospitals. Linear regression, decision tree, random forest and naïve bayes algorithms are used for estimation based on symptoms. Web framework is built on flask and overall system has accuracy of 88%.

As [22] mentioned paper about Symptom based health status prediction via Decision tree XGBoost, LDA SVM and random forest" published on March 2023 by Elif Meric, Cagdas Ozer deals with decision making activities for health status prediction using multiple models with multiple symptoms input along with their evaluation and it has increased the accuracy of model up to 98%. According to result of this project random forest is most efficient among mentioned algorithm

IV. PROPOSED SYSTEM

With a solution of providing multiple disease prediction based on variety of data about symptoms and medical record proposed system is integration of web development and machine learning technology. Collectively It includes majorly three use cases i.e. for predicting disease, for contact with medical experts and tips regarding health. Proposed system is web enabled that's why it is easily accessible from anywhere anytime. It can be a first choice of person to clear their doubts and take remedial measures based on current medical situations before appointing to doctors. It can accept inputs as symptoms as well as medical records and reflects the possibility

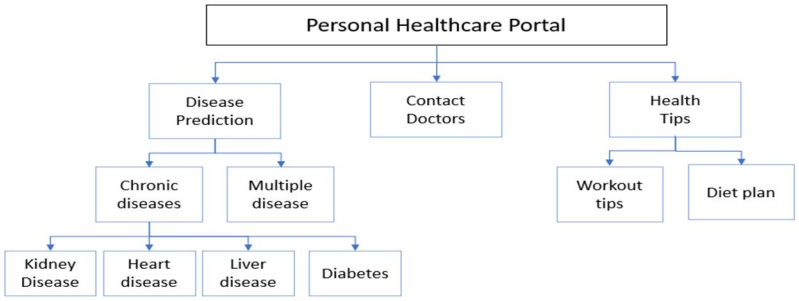


Fig 4.1: Block diagram of personal healthcare portal

Detailed explanation about modules in Fig 3.1 is given below-

1) *Disease Prediction*: It is the core module in this project which deals with two types of disease predictions i.e. multiple disease prediction and other is chronic disease predictions. Multiple disease prediction section takes inputs of symptoms from users and predict disease which can cause risk. Whereas chronic disease section comprises of 4 different separate disease prediction. It simply takes inputs of medical records of user and predict that the specific disease test is positive or not.
2) *Contact doctors*: this section is complementary for these who need to communicate with medical expert. It simply has list of various doctors along with their names, address and specialty. According to requirement user can see the contact details here and also can search nearby doctor by filtering out with entering pin code.
3) *Health tips*: This referred module is useful for information sharing about health consciousness by guiding users about workout, diet plans etc. here user can see different sections like yoga, exercises and nutrition. User will able to see that information directed by specific section

A. Algorithms Used

1) *Support Vector Machine:* Support Vector Machine is a supervised machine learning algorithm that which is used for both classification and regression. Each data item is plotted as a point in n-dimensional space where n denotes the number of features that are present. The value of each feature is the value of a particular coordinate on that dimension. After that, we find a hyper plane that will differentiate the two classes to perform classification. There can be multiple hyperplanes, we need to select the one that segregates the two classes better [14].
2) *Random Forest classifier:* Random Forest is a supervised machine learning algorithm that which is used for both classification and regression. Instead of using one model, it uses an ensemble of decision trees. All the different models are trained independently. The models are rained on different subsets of data. Some models may predict the correct result while others may not. After the results from all the models are taken, the final output is based on majority voting. This helps in predicting the correct outcome [14].
3) *Decision Tree Classifier:* Decision Tree is a supervised machine learning algorithm that which is used for both classification and regression. A decision tree is a flowchart-like structure in which each node represents a decision on a feature, each leaf node represents a class label or a decision taken after computing all features. Branches represent conjunctions or possible values of the features thatlead to those class labels [14] .

V. IMPLEMENTATION AND RESULTS

A. Implementation Methodology

As discussed earlier about design of the system for meeting the requirements of any machine learning models about disease prediction a suitable relevant data is necessary train the algorithm. The whole process of implementation followed similar approach of any machine learning process from data collection, data cleaning, exploratory data analysis, training and testing data models and evaluation. Dataset for various chronic diseases and multiple diseases with symptoms are collected from Kaggle. After getting structured data, data preprocessing is carried out for data cleaning like managing empty and undesirable data items. All resources for implementation are gathered and then arranged in suitable manner. After standardization of data Exploratory Data Analysis is performed for gaining more detailed view about dataset and on summarizing that feature engineering tasks carried out to make algorithm more efficient [17].

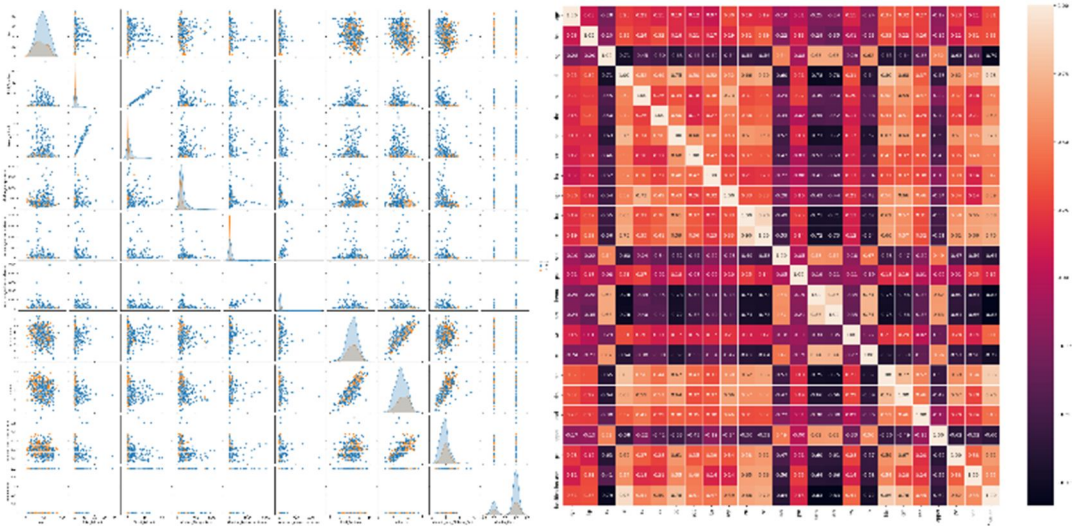


Fig 5.1.1: EDA (correlation estimation by pair-plot & heatmap)

Proposed system deals with user interaction with help of GUI. Web development frameworks like ReactJS are used for that purpose to create front end of portal. Also, modules like health tip suggestion is done successfully by connecting the system interface with desired APIs. REST API technology is used for directing user to the health-related information. Doctor's section developed by collecting doctor's information from trusted resources and centers and that data is organized in user friendly manner. The entire system is integrated with frameworks, APIs and different sections. For machine learning implementation website made suitable for taking inputs and able to reflecting outputs. UI design is made simple and attractive [12].
Flow chart for system is shown below:

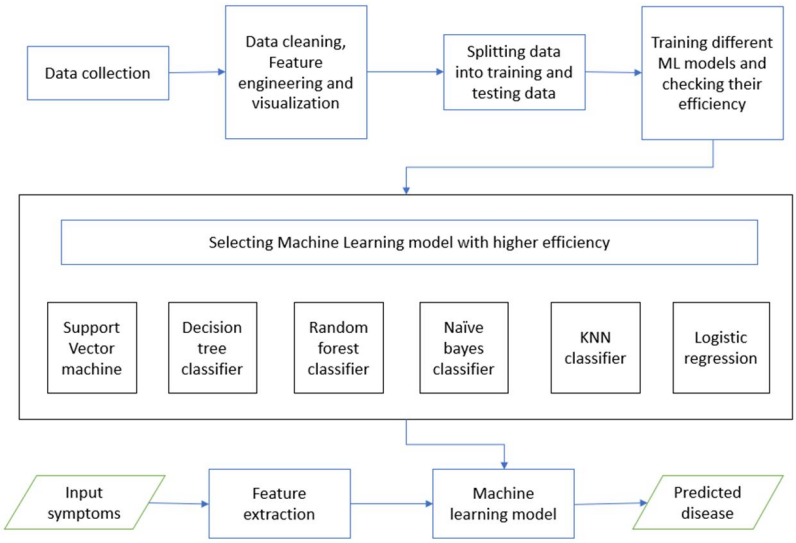


Fig 5.1.2: Flow chart for working of Personal healthcare portal

After developing framework of portal complementary with machine learning system, refined dataset is split for training and testing of various machine learning models. Majorly SVM, KNN, Random forest, Decision tree and logistic regression algorithms are implemented and then simultaneously compared & evaluated their performance[4]. Best suited model with high accuracy is used for prediction and that is connected with developed web portal.
The entire project implementation is carried out by waterfall model. Finally, it get tested to check for desired outputs. And Successful implementation of project is approved by respective authorities.

B. Results

Various Machine learning models are evaluated based on their accuracy and most efficient model is selected amongst them.

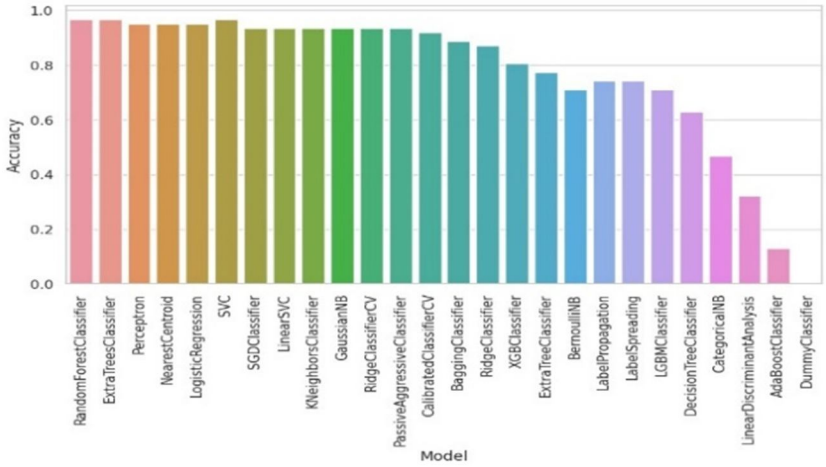


Fig 5.2.1: Comparison of ML models for multiple disease prediction

For Multiple disease prediction Random Forest classifier model is used due to its high accuracy about 97% and f1 score is approximately 0.9677.

SN.	Models	Accuracy	F1- score
1	SVM	0.96	0.96
2	Random forest	0.97	0.97
3	Decision tree	0.82	0.81
4	KNN	0.94	0.93
5	Logistic regression	0.95	0.95
6	Naïve bayes	0.93	0.94
7	XGBoost	0.80	0.81

Table 5.2.1: performance measures of Multiple disease prediction module

Confusion matrix of every used ML model is derived to check its other performance measures like precision, recall, sensitivity etc. Number of true positives, true negatives, false positive and false negative predicted and actual values can be estimated to check efficiency of any model as shown above.

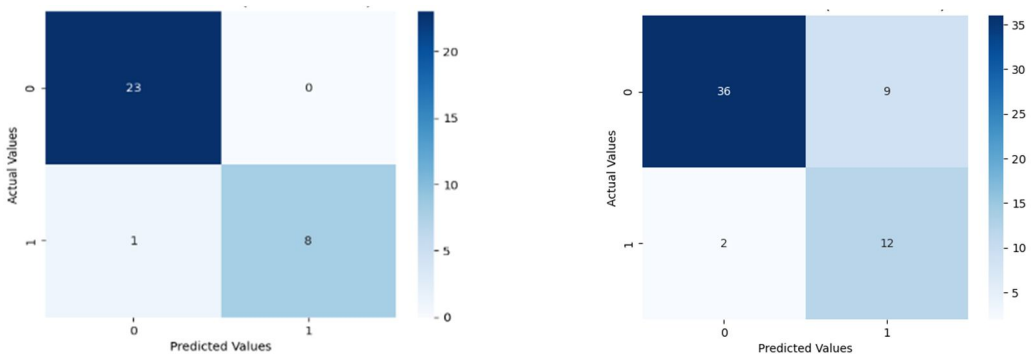


Fig 5.2.2: confusion matrix for kidney and liver disease prediction

Effectiveness of any ml model can be estimated by some standard scores like accuracy, F-1 score, ROC-AUC and some others [10]. Based on that algorithms are selected having best values of that measures taking less time.

SN.	Models	Accuracy	ROC AUC	F1- score
1	SVM	0.80	0.80	0.80
2	Random forest	0.84	0.82	0.83
3	Decision tree	0.79	0.78	0.79
4	KNN	0.82	0.82	0.82
5	Logistic regression	0.79	0.79	0.79
6	Naïve bayes	0.82	0.83	0.82
7	XGBoost	0.77	0.77	0.76

Table 5.2.2: performance measures of ML models for heart disease prediction

After evaluating various models regarding chronic disease predictions we have used Random Forest for heart disease prediction and Decision tree for kidney disease prediction. XGBoost and SVM classifiers are used for liver disease and diabetes prediction respectively.

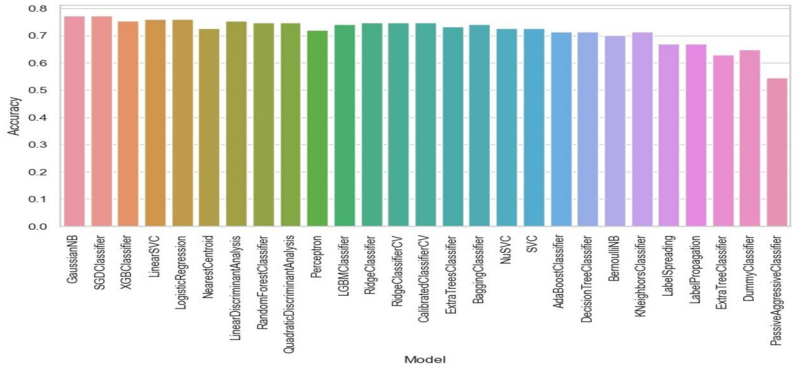


Fig 5.2.3: Comparison of ML models for diabetes prediction

According to mentioned study and results accuracy measure of heart disease prediction is 84%, kidney disease prediction is 98% whereas diabetes prediction has 80% and liver disease prediction has 83% accuracy.

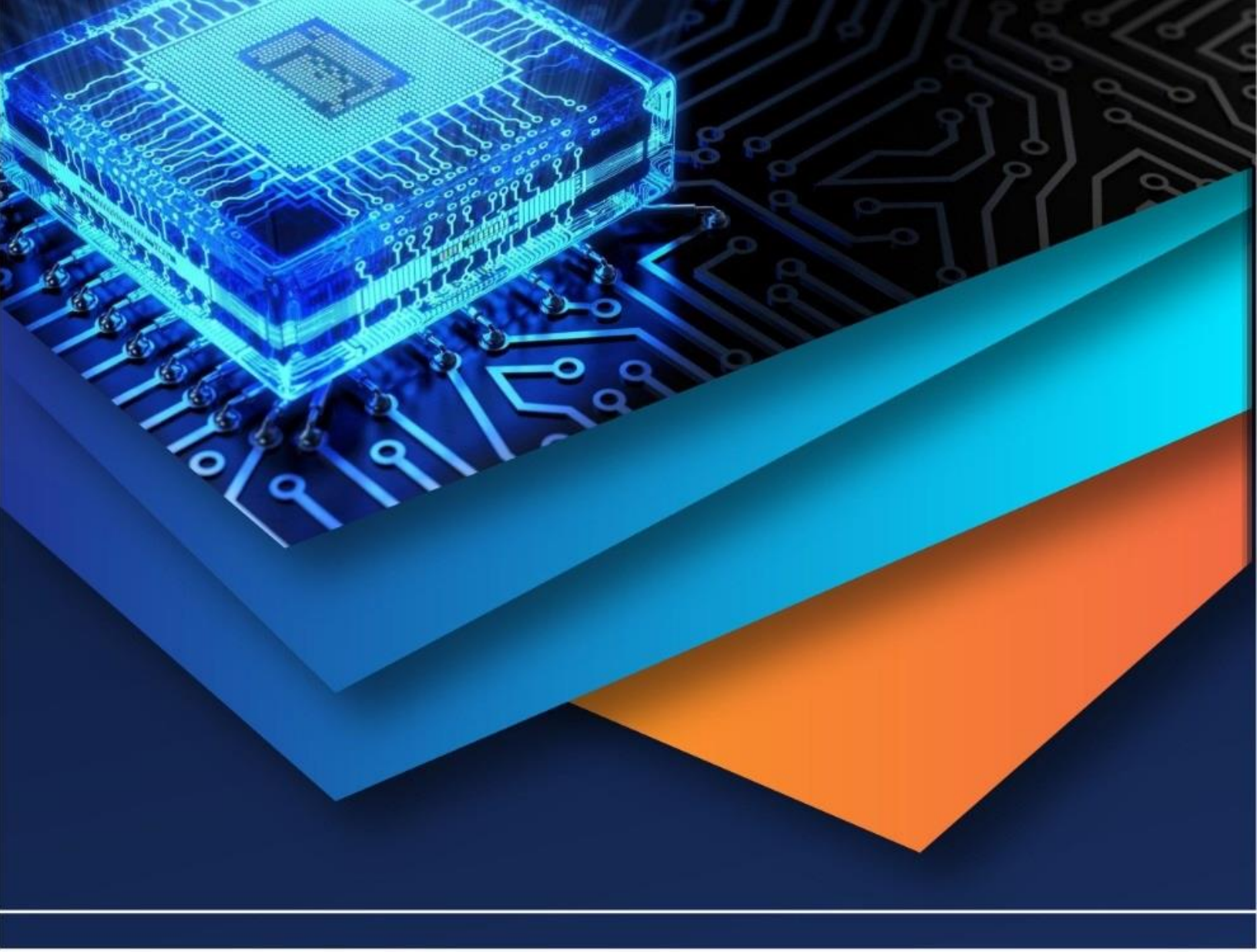
All algorithms are integrated with web- portal which is majorly related to interface interaction and information sharing to users. Well-designed navigation bars, search boxes are created for better experience. Textbox to take inputs for chronic disease prediction and checkbox for selecting inputs from list of symptoms for multiple disease prediction are there. Various clickstreams and connections with APIs are provided for directing to information resources about yoga, exercise and diet. Overall execution of ML model in web based framework is performing in proper way

VI. CONCLUSION AND FUTURE SCOPE

Provided system majorly deals with multiple disease prediction as well as specific chronic disease prediction by comparative analysis of various ML algorithms. We can conclude that entire system is built as per desired way with overall accuracy around 90%. Also API integration and connection with doctors details is done successfully. Proposed system may contribute much more in healthcare industry in some future advancements. Accuracy of the model can be increased by finetuning the hyperparameters and increasing size of datasets, using ensemble learning and deep learning algorithms. In future more changes can be done like connecting with medical centers as fully developed online medical service platform. This model could be integrated along with various recommendation systems, database and shopping-delivery of pharmaceuticals/medicines and also with government as well as private medical systems.

REFERENCES

[1] Sabyasachi dash, Sandeep kaushik, " Big data in healthcare: management, analysis and future prospects, June 19,2019.
[2] Touqeer Ahmed, Saeed Mian Qaisar, "Novel Web based multi class heart disease prediction using machine learning" March 2023, International conference of electrical engineering and sustainable technologies (ICEEST).
[3] Abhishek Singh, Dr SPS Chauhan, "Prediction and Analysis of Multiple diseases using Machine learning Techniques" Volume:07 issue: 04, April 2020, international journal of engineering and technology, Greater Noida, India e-ISSN : 2395-0056.
[4] Shahadat Uddin, Arif Khan, Mohammad ali-Moni, "Comparing different supervised machine learning algorithms for disease predictions", December 2019, BMC Informatics and decision making, DOI: :10.1186/s12911-019-1004-8.
[5] Harish Kumar Rajora, Narinder singh Punn, Sonali Agarwal, "Web based disease prediction and recommender system", June 2021, Indian Institute of Information Technology, DOI:10.1007/978-981-19-0840-8_23.
[6] Marouane Fethi Ferjani, "Disease Prediction using Machine learning", December 2020, Computing Department of Bournemouth University, England, DOI:10.13140/RG.2.2.18279.47521.
[7] Nandikatla Chandrasekharan, Samineni Peddakrishna, "Enhancing heart disease prediction accuracy through machine learning techniques and optimization", April 2023, India, DOI:10.3390/pr11041210.
[8] Arumugam K, Priyanka Shinde, Mohd Naved, Tatiana Gonzales "multiple disease prediction using ML algorithms", August 2021, DOI: 10.1016/matpr.2021.07.361, scientific committee of international conference on nanoelectronics, nanomaterials & nanotechnology.
[9] Nikhila," Chronic disease prediction using machine learning ensemble algorithm", April 2021, Published in international conference on computing, communication and intelligence system, DOI: 10.1109/ICCCIS51004.2021.9397144.
[10] Amit Juyal, Chetan Pandey, Janmejay pant, Dr Ankur Dhumka, "Performance analysis of supervised Machine learning on medical dataset", March2020, DOI:10.35940/ijrte.F7908.03862067.
[11] Prof. Archana Ugale, Abhijeet Gadakh, Roshan Sawant, "Integrated healthcare system using ML", November 2022, DOI:10.22214/ijraset.2022.47520.
[12] Adnan Qayyum , Junaid Qadir , Muhammad Bilal , and Ala Al-Fuqaha, Secure and " Robust Machine Learning application for Healthcare : A Survey" https://ieeexplore.ieee.org/iel7/4664312/9333686/09153891.pdf, 2021.
[13] M. McDermott, S. Wang, N. Marinsek, R. Ranganath, M Ghassemi, and L. Foschini, "Reproducibility in machine learning for health," Presented at Internation Conf. Learn. Representative Reproducibility Mach. Learn. Workshop, 2019.
[14] Alloghani D. Al-Jumeily, J. Mustafina, A. Hussain, and A. J. Aljaaf, "A systematic review on supervised and unsupervised machine learning algorithms for data science," in Supervised and Unsupervised Learning for Data Science. Berlin, Germany: Springer, 2020, pp. 3–21.
[15] R Sykes, Syed Tanbeer, "MyHealthPortal – A web-based e-Healthcare web portal for out-of-hospital patientcare" https://journals.sagepub.com/doi/full/10.1177/2055207621989194, February 24,2019.
[16] Jayram Ramesh , Raafat Aburukba, Assim Sagahyrun, "A remote healthcare monitoring framework for diabetes prediction using machine learning", may 2, 2021.
[17] Ajinkya Kunjir, Harshal Sawant, 'Data mining and visualization for prediction of multiple diseases I healthcare', 2017 IEEE.
[18] Dhiraj Dahiwade, Prof. Gajanan Patle, Prof. Ektaa Meshram "Designing Disease Prediction Model Using Machine Learning Approach" ©2019 IEEE.
[19] Poojitha Amin ,Nikhitha R. Anikireddypally ,Suraj Khurana, "Personalized Health Monitoring Using PredictiveAnalytics" DOI:10.1109?BigDataServices.2019.00048 9787288100593/19$31.00©2019. IEEE.
[20] Stepheny Lucas, Mitali Desai, Amisha Khot, Sincee Harriet, Nilambari Narkar, "Smartcare: A Symptom based disease prediction system using Machine Learning approach", December 11 2022, DOI : 10.22214/ijraset.2022.47434
[21] Farukh Aslam khan, khan Zeb, Mabrook al-Rakhami, "Detection and Prediction of diabetes using data mining: a comprehensive review" February 2021, Saudi Arabia, DOI : 10.1109/ACCESS.2021.3059343. IEEE.
[22] Elif Meric, Cagdas Ozer, "Symptom based health status prediction via Decision tree XGBoost, LDA SVM and random forest", March 2023, DOI:10.1007/978-3-031-27099-4_15



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY