International Journal of Recent Advances in Multidisciplinary Topics
Volume 2, Issue 10, October 2021
https://www.ijramt.com | ISSN (Online): 2582-7839

Multiple Disease Prediction System: A Review

Aayushi Agrawal[1*], Shubham Dalal[2], Dhvanan Rangrej[3], Nitin Choubey[4]
[1,2,3]Student, Department of Information Technology, SVKM's NMIMS Mukesh Patel School of Technology Management and Engineering, Shirpur, India
[4]Head of the Department of Information Technology, SVKM's NMIMS Mukesh Patel School of Technology Management and Engineering, Shirpur, India

***Abstract*: Machine Learning techniques are used for a lot of applications. In healthcare, machine learning plays a critical role in disease prediction. For detecting a disease, several tests must be required from the patient. By using machine learning techniques, the number of tests can be reduced. This reduced test performs a critical function in time and performance. This article analyzes machine learning strategies that can be used to predict distinct varieties of diseases. This paper reviewed the research papers which especially deal with predicting Diabetes, Heart disease, and Breast cancer. This article presents a review of various models based on such algorithms, techniques, and an analysis of their performance. Research has been carried out on various models of supervised learning algorithms and some of them are Support Vector Machine (SVM), K Nearest Neighbor (KNN), Decision Tree, Naïve Bayes, Convolutional Neural Network and Random Forest (RF).**

***Keywords*: Support Vector Machine (SVM), Diabetes, Naive Bayes, Heart Disease, Breast Cancer, Convolutional neural network (CNN), Data Mining, K Nearest Neighbor (KNN), Decision Tree , Random Forest.**

1. Introduction

Multiple Disease Prediction using Machine Learning is a machine that predicts the sickness primarily based totally at the facts supplied via way of means of the user. It predicts the disease of the patient, or the user based on the information, or the symptoms entered by the user into the system and provides accurate results based on that information. Nowadays health industry plays a major role in curing the diseases of the patients so this will be a help for the health industry to inform the user and it will be useful for the user in case they don't want to go to the hospital or any other clinics. This system acts as a recommendation system to doctors and the health industry for the treatment of the patient and further actions regarding the treatment can be solely concluded by the doctors.

2. Process

The system is designed in such a way that the user will find no difficulty while operating the system. So, when a user accesses the website, they will have to register themself. They will then need to log into the web application. The credentials will be authenticated in the backend. If the credentials are not valid then they will be redirected to the same page and if those credentials are valid then they will be directed to the homepage. On the homepage, they get to choose their disease about which they want to know. Once they select the disease, they will be redirected to the input page where they will have to provide the valid details that are asked. For their understanding, they will be provided with the format of the input. After this, the values will be sent to the model and the model will run these values through their database and provide with most accurate output. During this process, if the values are invalid then a decision will be taken where the user will be redirected to the same page. If the values are valid then there will be a prediction in the form of output. The prediction will show whether the user has been affected by the disease or not.

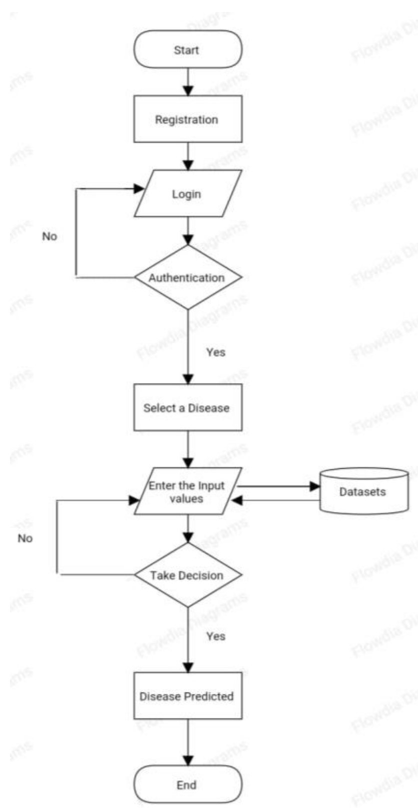

Fig. 1. Multiple Disease Prediction System

*Corresponding author: aayushia2@gmail.com

3. Literature Review Analysis

1) Predicting Diabetes in Healthy Population through Machine Learning

This research paper throws light upon the model which predicts the high-risk population so that a timely population-based intervention could prevent future complications. Researchers revisit the information of the San Antonio metropolis Heart Study and use machine learning to predict the long-term development of type-2 polygenic disease. Researchers used the linear support vector machines to construct a prediction model of future development of type-2 diabetes. To construct the prediction model, they use the help vector machines and ten capabilities which can be widely recognized withinside the literature as sturdy predictors of destiny diabetes. Development of the classifiers on an unbalanced dataset poses an average machine learning problem that outcomes with inside the educated fashions being biased toward the majority class. In this study, they balanced the two classes with the aim to get unbiased models in the training. The classification threshold which controls the probability that a sample belongs to a certain class can be modified to maximize the true positive rate (recall) of the classifier. The results of this study show a validation accuracy of 84.1% with a recall rate of 81.1% averaged over 100 iterations. The outcomes of the study show that high values of glucose observed at the 2-hour mark during the oral glucose Tolerance testing (OGTT) can strongly indicate the potential risk of future development of type 2 diabetes [1].

2) Applying Machine learning methods in Diagnosing Heart disease for Diabetic Patients

This research article sheds light on a model that has suggested a method to demonstrate that mining may assist to retrieve relevant correlation even from features that are not direct indicators of the class they are attempting to predict. In their study, they attempted to predict the probability of developing a heart disease using attributes derived from diabetes diagnosis, and they demonstrated that it is feasible to identify heart disease susceptibility in diabetic patients with fair accuracy. This type of classifier can aid in the early detection of a diabetic patient's sensitivity to heart disease. Patients is probably cautioned to adjust their life-style due to this. This will prevent diabetes people from developing heart disease, resulting in lower death rates and lower healthcare costs for the state. In this study, SVMs were shown to be a supervised learning model with high predictive performance and were also investigated using the ROC curve for training and test data. Therefore, they proposed the SVM model to categorize the diabetes data set [2].

3) Designing Disease Prediction Model Using Machine Learning Approach

People face numerous diseases because of the condition and their living habits. That the prediction of sickness at an associate degree earlier stage becomes a crucial task. However, the correct prediction of the idea of symptoms becomes too problematic for physicians. The most difficult task is the correct prediction of the disease. With the assistance of sickness knowledge, data processing finds hidden pattern data within the large quantity of medical knowledge. Researchers projected general sickness prediction supported symptoms of the patient. For sickness prediction, they use K-Nearest Neighbor (KNN) and Convolutional neural network (CNN) machine learning algorithms for correct prediction of sickness. For sickness prediction needed sickness symptoms dataset. During this general sickness prediction, the living habits of someone and examination info area unit thought of for the correct prediction. The accuracy of general sickness prediction by exploitation CNN is 84.5% that is over KNN formula. They planned a general illness prediction system that supported machine learning algorithms. They used KNN and CNN algorithms to classify patient knowledge because of it must method existing knowledge for predicting actual illness supported symptoms. They got correct general illness risk prediction as output, by giving the input as a patient record that helps them to know the extent of illness risk prediction. Attributable to this method could cause low time consumption and stripped price attainable for illness prediction and risk prediction. Researchers compare the results between KNN and CNN algorithmic rule in terms of accuracy and time and therefore the accuracy of CNN algorithmic rule that is over KNN algorithmic rule and time needed for classification for CNN is a smaller amount than KNN. So, we can say CNN is better than KNN in terms of accuracy and time. [3]

4) Breast Cancer Detection and Prediction using Machine Learning

In this studies paper, we've discovered to construct a breast most cancers tumor predictor at the Wisconsin dataset and steps to create graphs and consequences for the same. It has been determined that a very good dataset affords higher accuracy. The choice of suitable algorithms with a very good domestic dataset can result in the improvement of higher prediction structures. These structures can help in the right remedy techniques for an affected person recognized with breast most cancer. There are many remedies for an affected person primarily based totally on breast most cancers stage; information mining and system studying may be an excellent assist in identifying the road of remedy to be observed with the aid of using extracting understanding from such appropriate databases. [4]

5) Breast Cancer Classification and Prediction using Machine Learning

This paper compares six machine learning (ML) algorithms on the Wisconsin Diagnostic Breast Cancer (WDBC) dataset, the dataset was divided into two phases for the implementation of the ML algorithms: training and testing. The algorithm that produces the best results will be utilized as the website's backend, and the model will then categorize cancer as benign or malignant. Machine learning algorithms may be applied in medical research, which enhances the field [5].

6) Machine Learning Implementation on Medical Domain to Identify Disease Insights using TMS

Health could be a person’s true wealth. Medicine is the biggest domain within the world that is providing a great deal of data for the researchers to implement one thing new from it. During this situation researchers are implementing machine learning methodology in health care data to know the human disease patterns and mistreatment that random forest and different machine learning models work and predict the procedure an individual has got to follow to urge a decent health and avoid the various health loss activities we tend to do often. They used TMS algorithmic rules to implement trust. Then they combined TMS algorithm with the ancient methodology algorithm to get this final algorithm:

TMS_Final (Variable, Ancient, Modern) {Disease X; Variables1 = Ancient (Variables); Variables2 = Modern (Variables); TMS(Variables1, Variables 2) Result(0) }

The result of using the algorithms is Random Forest acquired the highest accuracy of the implementation. Using the modern mechanism in the medical they were not getting the accurate results and the implementation lies in combining them both. Because ancient mechanisms are the low side effect things and the implementation after combining them will give the most accurate results in the implementation of the treatments. The treatment using this kind of mechanism will be helpful for better results [6]. Using Electronic Health Records and Machine Learning to Make Medical Related Predictions from Non-Medical Data: This paper focuses on testing the hypothesis that the application of machine learning techniques on data (Administrative HIS (Hospital Information System) and EHR (Electronic Health Record)) of this nature can be used to predict problems in the Health IT domain. Administrative HIS (Hospital info System) and EHR (Electronic Health Record) information area unit characterized by lower privacy sensitivity, therefore easier movability, and handling, in addition as higher info quality. They experiment with the prediction of the probability of early readmission at the time of a patient's discharge. They extract real HIS data and perform data processing techniques. Then they apply a series of machine learning algorithms and measure the performance of the emergent models. They acknowledge many limitations associated with each approach and selected methodology: the very fact that the information origin is merely one hospital implies that attainable admissions in different hospitals don't seem to be accounted for, and perhaps introduces a bias associated with the policies or the quality of care of the precise establishment. Results: All applied strategies performed well on top of random guesswork, even with lowest hyper-parameter standardization. Thus, provided that the experiments give proof in favor of the underlying hypothesis, future experimentation on additional fine-tuned (thus additional robust) models may lead to applications fitted to productive environments [7].

7) Multi Disease Prediction using Data Mining Techniques

This research article examines data mining approaches that can be used to forecast various sorts of illnesses. This study examined research studies that mostly focused on predicting heart disease, diabetes, and breast cancer, among other things. In this work, two distinct data mining classification approaches were used to predict various diseases and their performance against choosing the best classifier. Building accurate and computationally efficient classifiers for medical applications is a significant problem in data mining and machine learning. article, two distinct data mining classification approaches were applied to predict various diseases and their performance in determining which classifier was better. Creating accurate and efficient classifiers for medical applications is a significant challenge in data mining and machine learning. [8]

8) Machine Learning based Medical Information Analysis, Estimations and Approximations over Present Health Research Domain

In this article, the researchers focus on the mechanism for accomplishing things like machine learning and artificial intelligence with healthcare as a single platform. The devices facilitate in creating the higher choices supporting the past collected data by the assistance of the intelligent process mechanism analyzing as well as big data techniques. The following things which are used for developing the work such as:

1. *Data Set:* The data used by researchers for the work is collected by the National Institute of Diabetes and Digestive and Kidney Diseases. The data collected was actual data collected from patients.
2. *Parameters:* The following are the terms and conditions which are to be mentioned for developing of the project such as:

- Disease information of patient
- Patient basic information such as name, place, gender etc.
- Collecting the blood and gathering the reports of blood tests for sugar, sample Blood Pressure, Plasma levels in body, Chromosomes levels, Skin health condition, Insulin condition and BMI.

The results we can see that they got are the results with an accuracy of 97.13% of efficiency. The input of the values based upon the following parameters namely: sugar, Blood Pressure, Plasma levels in body, Chromosomes levels, Skin health condition, Insulin condition, BMI, and Outcome. This paper offers a summed-up structure to personalized social welfare that impacts the upsides of remote observant, distributed computing, huge data, and responsive AI. It offers organized thanks to influence to facilitate quickly developing data of people with extreme ailments however it wants complicated programming associated with the employment of an assortment of Refine, KNIME and knowledge mining to be utilized to enhance the proficiency of knowledge examination.[9]

9) Prediction of heart disease using a hybrid technique in data mining classification

Predicting cardiopathy is considered the foremost sophisticated task within the field of bioscience, hence the need to develop a decision support system to detect heart disease in a patient. We propose an efficient hybrid genetic algorithm that uses the back-propagation technique. Approach to predicting heart disease. Nowadays medicine has come a long way to treat patients with different types of diseases. Among the most threatening are heart diseases that cannot be seen with the naked eye and appear immediately when their limits are reached. Bad clinical decisions would result in the death of a patient, which no hospital can handle. To achieve correct and inexpensive treatment, computer and support systems can be developed to help make good decisions. Many hospitals use hospital information systems to manage their medical care or patient records. These systems produce large amounts of data in the form of images, text, graphics, and numbers. Unfortunately, this knowledge isn't wont to aid medical decision-making. There is much information hidden in this data that has yet to be explored, leading to an important question of how to turn data into useful information. So, there is a need to develop a great project that will help clinicians predict heart disease. The main aim of this work is to develop a prototype that can determine and extract unknown knowledge (patterns and relationships) associated with heart disease from a database of preceding heart diseases. Help clinicians make smart medical selections that conventional selection help structures couldn't. Efficient remedies will let you lessen remedy costs [10].

10) HDPS, Heart Disease Prediction System

The prognosis of heart disease relies upon on a complicated mixture of scientific and pathological information in maximum of cases. Because of this complexity, there's a remarkable hobby amongst the ones professional with inside the artwork and scientific researchers in green and correct prediction of coronary heart ailment. We have developed a heart disease prediction system that can help clinicians predict the status of heart disease based on clinical patient data. Our approaches involve three steps: First, they chose thirteen key clinical traits i.e., age, gender, pain type, trestbps, Cholesterol, fast blood sugar, resting ECG, most heart rate, exercise-induced angina, previous peak, steepness, variety of colored vessels and the like. Second, they developed a man-made neural network algorithmic rule to classify heart diseases supported these clinical options. The prediction accuracy is almost 80%. The HDPS system will consist of several functions including the clinical data entry section, the ROC curve display section, and the predictive performance display section (run time, precision, sensitivity, specificity, and prediction of the result). Heart disease of the patient. The HDPS system developed during this study may be a novel approach that may be utilized in the classification of heart condition [11].

11) Disease Risk Prediction by Using Convolutional Neural Network

Data analysis plays an important role in dealing with a large amount of data in healthcare. Previous medical research has relied on the processing and assimilation of a large amount of hospital data rather than prediction. Due to a tremendously growing quantity of information within the medicine and healthcare fields. Accurate analysis of medical data contributes to earlier disease detection and patient care; however, the accuracy decreases if some medical information is missing. To beat the matter of missing medical data, we tend to perform data cleansing and imputation to rework incomplete data. We are performing on the prediction of heart disease based on the information set using Naive Bayes and therefore the ANN rule. To increase this work, we propose predicting disease risk with structured data. We tend to use a risk prediction algorithm for unimodal diseases based on convolutional neural networks. The prediction accuracy of the CNNUDRP algorithm reaches quite 65%. In addition, this system answers the question about the disease that people face in their life [12].

12) Comparison of Machine Learning Algorithms for Prediction of Diabetes

The machine learning method is increasingly used for data analysis where large databases need to be analyzed. one in all the areas wherever solutions embrace the employment of machine learning is in medical prediction, which is employed to seem at the chance that an individual could have an illness within the future. One in all the areas of medical prediction within which machine learning solutions are used is predicting diabetes. Polygenic disorder could be a disease that's more and more gift today. This text presents a comparison of the experimental results obtained exploitation 3 machine learning algorithms to predict diabetes. The three algorithms thought about are Support Vector Machine, Naive Bayes, and Random Forest. The target of this work is to research the performance of the algorithms considering totally different metrics to check different techniques to get larger precision. We tend to find that Support Vector Machine and Random Forest achieved quite 80% accuracy. [13]

13) Diabetes disease prediction using data mining

Data mining is a branch of software engineering. It is the methodological system used to discover examples in large information sets, such as strategies on the intersection of manufactured intelligence, machine learning, knowledge, and database systems. The aim of the data mining method is to consider information from a dataset and convert it right into a significant shape for later use. Our research focuses on this part of the medical reasoning learning design through data collected from diabetes and creating smart therapeutic options that form an emotional support network to help clinicians. The main objective of this study is to create an intelligent diabetic disease prediction system that allows the analysis of diabetic disease using the database of diabetic patients. In this system, we propose the use of algorithms such as Bayesian and KNN (K-Nearest Neighbor) to collect and analyze the database of diabetic patients using various attributes of diabetes to predict diabetic diseases [14].

14) A novel breast cancer prediction system

This article presents a novel design and methodology implemented through a web-based decision support system developed for oncologists and healthcare professionals. Provides an overview of the development process and methodology. It also examines previous work and similar research in this area, highlighting some of the limitations in these systems and comparing our system, which addressed these shortcomings, as well as new features that should be considered for future work.

The study shows that data mining and knowledge discovery technology can be used to develop computer models that uncover hidden patterns and useful information from large data sets. In healthcare, technology can examine 4.444 billion medical records. Computer models are used by modern physicians to aid in their decision-making, diagnoses, and treatments. Oncologists and healthcare professionals will also need to analyze the results of the system. Our additional work would include the additional inclusion of medical variables such as chemotherapy and gene therapy to add to the prediction data sets and add advanced functionality to the prototype system [15].

15) Comparative Study of Machine Learning Algorithms for Breast Cancer Prediction

Breast cancer is the most common cancer in women and an estimated 270,000 new cases were diagnosed in 2019. Because of this, detection software is needed to identify it before it becomes fatal. Using machine learning algorithms, software can be developed that detects and treats this dangerous cancer before it can lead to the death of the patient. It is also the most common cancer among Indian women and a woman with breast cancer has a 50% chance of survival. Many machine learning algorithms can be used for breast cancer detection.

This article proposes two machine learning algorithms to compare the logistic regression and decision tree algorithms in the Wisconsin data set (diagnosis) and use the algorithm with the highest precision to predict breast cancer. In this research, two machine learning algorithms, the decision tree classifier and logistic regression, are implemented to predict breast cancer and the accuracies of both are compared to find which of the two is better at predicting. The decision tree classifier is the most suitable algorithm for prediction, because it had precise prediction precision by using it [in the "Wisconsin Breast Cancer (Diagnostic) Data Set"]. Given the characteristics of this data set, breast cancer can be predicted with near millimeter precision using our Decision Tree Classifier algorithm [16].

16) Intelligent Breast Cancer Prediction Model Using Data Mining Techniques

The incidence of breast cancer has continuing to rise for the past twenty years. So, the identification and treatment of breast cancer became a task of utmost urgency. During this study, we tend to shall use data processing techniques to make a diagnostic model for breast cancer. The results of the experiments show that the employment of the feature choice technique greatly improves the accuracy of the diagnostic model and at the same time selects nine features relevant and important for the diagnosis of breast cancer. The breast cancer diagnostic model constructed in this study is easy to generalize. In this study, our goal is to build a diagnostic model for breast cancer that will examine the relationship between breast cancer and its symptoms.

A feature selection process, INTERACT, is used to select related and important features to improve the accuracy of the diagnostic model. And SVM is applied to build the classification model. To exhibit the significance of characteristic choice, diagnostic models with and without characteristic choice are created. The experiments significantly improve the accuracy of the feature selection diagnostic model compared to the non-feature selection model. Meanwhile, nine characteristics have been selected as relevant factors for the construction of the diagnostic model. The information we find in this study can be complementary information for the treating physician who can better diagnose heart disease [17].

17) Machine Learning Models in Type 2 Diabetes Risk Prediction: Results from a Cross-sectional Retrospective Study in Chinese Adults

This has improved diabetes prognosis and estimate among many groups of people. The authors collaborated to create a new approach for spotting undiagnosed models using data from a hospital. The answer was discovered by analyzing the information of 8000 non-diabetic individuals and 3845 diabetic patients. The AUC of the five models combined is 0.97. These algorithms have the highest prediction accuracy (ACC = 0.8084) with random forest. As a result, the prediction of early patients using the AUC model is in better shape in the health industry. The research article presents the significant results of merging five models applying machine learning components and offers the accuracy rate for comparison. The usage of five models has also been highlighted in many research papers, providing a certain degree of accuracy [18].

18) Machine Learning for Health Services Researchers

This paper expands on the understanding of the dataset utilized for illness prediction. They concentrate on utilizing the data set in an instructive manner rather than in a setting of poor quality. The study focuses on the use of a machine learning model to predict the outcome. It additionally stresses the negative components of prediction and therefore the consequences of incorrect prediction. This study gave a superior account the machine learning model. Machine learning could be tough to implement in some industries. The approach gives a comprehensive knowledge of the dataset and its practical usage, while minimizing prediction error [19].

19) Ensemble Learning on Diabetes Data Set and Early Diabetes Prediction

Joint learning is a process of using multiple classifiers to find a solution to a problem. Using joint learning on a diabetes data set, a model is trained to predict the early onset of diabetes. This model aims at two main goals: whether the person is at risk for diabetes in the near future and the risk of developing diabetes that is associated with the person. The data set consists of a basic set of questionnaires that could be useful for training the model based on various classifiers or experts strategically training the model. The libraries used in the system are Scikit Learn, Pandas, and NumPy. Once trained, the model does not require retraining and could be useful in predicting the early onset of diabetes in a person based on a basic lifestyle questionnaire, such as regular lifestyle habits, eating habits, family history, and medical history. Since predicting diabetes before it occurs is a set of problems that must be solved, since the disease is basically a lifestyle disease that requires patient self-assessment and disease prevention at every step, joint learning to the prediction is used to reduce the uncertainty of the risk of diabetes. Since ensemble learning is a collection of different machine learning models [20].

20) Prognosis of Diseases Using Machine Learning Algorithms, a Survey

With the rapid growth of data in biomedicine and healthcare, it has been a difficult task to quickly extract knowledge from the available data. Predicting disease can reduce the risk of life; many studies have shown the connection between machine learning algorithms in various disease predictions. Because healthcare needs timely and accurate performance for diagnosis, we use big data analytics to better understand disease stage, predict disease, make effective decisions, maintain electronic medical records, and organize structured and unstructured data to achieve accurate results. This document summarizes a traditional survey and the importance of big data analytics in healthcare for various applications in which machine learning algorithms are implicit [21].

21) Prognosis of Diseases Using Machine Learning Algorithms, a Survey

With the rapid growth of data in biomedicine and healthcare, it has been a difficult task to quickly extract knowledge from the available data. Predicting disease can reduce the risk of life; many studies have shown the connection between machine learning algorithms in various disease predictions. Because healthcare needs timely and accurate performance for diagnosis, we use big data analytics to better understand disease stage, predict disease, make effective decisions, maintain electronic medical records, and organize structured and unstructured data to achieve accurate results. This document summarizes a traditional survey and the importance of big data analytics in healthcare for various applications in which machine learning algorithms are implicit [20].

4. Conclusion

Finally, I'd wish to emphasize that this project Multiple disease prediction system using machine learning is extremely helpful in everyone' daily life, however it's particularly vital for the healthcare sector, because of they're the ones who use these systems daily to predict the diseases of patients based on their general information and the symptoms that they have. This is a web-application based project which is intended for doctors and co-workers in hospitals who can refer this system as a recommendation system to help their patients. It will also act as the connecting bridge between the model and the user to help perform a basic diagnosis. Through this project, we intend to aware people from the diseases that they may have, either earlier or semi-earlier stages. Also, this project will act as one stop location for understanding some common diseases.

References

[1] H. Abbas, L. Alic, M. Rios, M. Abdul-Ghani and K. Qaraqe, “Predicting Diabetes in Healthy Population through Machine Learning,” 2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS), 2019, pp. 567-570,

[2] G. Parthiban, S.K.Srivasta ”Applying Machine learning methods in Diagnosing Heart disease for Diabetic Patients” International Journal of Applied Information Systems(IJAIS) Foundation of Computer Science FCS, New York, USA vol. 3– no.7, August 2012.

[3] D. Dahiwade, G. Patle and E. Meshram, "Designing Disease Prediction Model Using Machine Learning Approach," 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), 2019, pp. 1211-1215,

[4] Chowdhury, Anoy. (2020). Breast Cancer Detection and Prediction using Machine Learning.

[5] Jean Sunny, Nikita Rane, Rucha Kanade, Sulochana Devi, 2020, Breast Cancer Classification and Prediction using Machine Learning, (IJERT) vol 09, no. 02 (February 2020)

[6] S. M. Sasubilli, A. Kumar and V. Dutt, "Machine Learning Implementation on Medical Domain to Identify Disease Insights using TMS," 2020 International Conference on Advances in Computing and Communication Engineering (ICACCE), 2020, pp. 1-4,

[7] S. Pitoglou, Y. Koumpouros and A. Anastasiou, "Using Electronic Health Records and Machine Learning to Make Medical-Related Predictions from Non-Medical Data," 2018 (iCMLDE),2018, pp.56-60,

[8] Kamaraj, K.Gomathi & Priyaa, D.Shanmuga. (2016). Multi Disease Prediction using Data Mining Techniques. International Journal of System and Software Engineering.

[9] R. Singh and N. Sharma, "Machine Learning based Medical Information Analysis, Estimations and Approximations over Present Health Research Domain," 2020 International Conference on Computational Performance Evaluation (ComPE), 2020, pp. 704-708,

[10] A. Dewan and M. Sharma, "Prediction of heart disease using a hybrid technique in data mining classification," 2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom), 2015, pp. 704-706.

[11] A. H. Chen, S. Y. Huang, P. S. Hong, C. H. Cheng and E. J. Lin, "HDPS: Heart disease prediction system," 2011 Computing in Cardiology, 2011, pp. 557-560.

[12] S. Ambekar and R. Phalnikar, "Disease Risk Prediction by Using Convolutional Neural Network," 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), 2018, pp. 1-5,

[13] N. E. Costea, E. V. Moisi and D. E. Popescu, "Comparison of Machine Learning Algorithms for Prediction of Diabetes," 2021 16th International Conference on Engineering of Modern Electric Systems (EMES), 2021, pp. 1-4.

[14] Shetty, K. Rit, S. Shaikh and N. Patil, "Diabetes disease prediction using data mining," 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), 2017, pp. 1-5.

[15] I. C. Tee and A. H. Gazala, "A novel breast cancer prediction system," 2011 International Symposium on Innovations in Intelligent Systems and Applications, 2011, pp. 621-625.

[16] P. P. Sengar, M. J. Gaikwad and A. S. Nagdive, "Comparative Study of Machine Learning Algorithms for Breast Cancer Prediction," 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), 2020, pp. 796-801.

[17] R. Shen, Y. Yang and F. Shao, "Intelligent Breast Cancer Prediction Model Using Data Mining Techniques," 2014 Sixth International Conference on Intelligent Human-Machine Systems and Cybernetics, 2014, pp. 384-387.

[18] X. L. Xiong, R. X. Zhang, Y. Bi, W. H. Zhou, Y. Yu, and D. L. Zhu, "Machine Learning Models in Type 2 Diabetes Risk Prediction: Results from a Cross-sectional Retrospective Study in Chinese Adults," Curr Med Sci, vol. 39, no. 4, pp. 582-588, Aug 2019.

[19] P. Doupe, J. Faghmous, and S. Basu, "Machine Learning for Health Services Researchers," Value Health, vol. 22, no. 7, pp. 808-815, Jul 2019.

[20] G. G. Warsi, S. Saini and K. Khatri, "Ensemble Learning on Diabetes Data Set and Early Diabetes Prediction," 2019 International Conference on Computing, Power and Communication Technologies (GUCON), 2019, pp. 182-187.

N. M. J. Kumari and K. K. V. Krishna, "Prognosis of Diseases Using Machine Learning Algorithms: A Survey," 2018 International Conference on Current Trends towards Converging Technologies (ICCTCT), 2018, pp. 1-9.

.