International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 10 Issue: 02 | Feb 2023 www.irjet.net p-ISSN: 2395-0072

Multiple Disease Prediction System: A Review

Urvaang Naik[1], Vipul Mashruwala[2], Kunj Joshi[3]

[1,2,3] *Student, Department of Computer Engineering, SVKM's NMIMS Mukesh Patel School of Technology Management and Engineering, Shirpur, Maharashtra, India*

Abstract*: Machine Learning and data mining technologies have revolutionized the healthcare industry by uncovering new, efficient ways to diagnose diseases. This research focuses on analyzing a variety of studies that employed predictive modelling for predicting heart disease, diabetes and breast cancer amongst other illnesses. Through employing an innovative Multi-Disease Prediction System based on users' given symptoms as inputted into the system, researchers are discovering effective methods in reducing time constraints while still achieving successful performance outcomes when diagnosing various medical issues. The Multi Disease Prediction System employs predictive modelling to identify diseases with incredible accuracy. By examining the user's symptoms, it can determine what medical conditions could be causing that particular set of signs and indicate the likelihood of each potential illness – delivering a powerful tool for detection straight into your hands.*

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

1. INTRODUCTION

Multiple Disease Prediction Using System Learning is a machine that predicts illness based only on information supplied by the user. Machine learning gives the medical field a strong platform for solving problems quickly and effectively. The healthcare industry is transforming with the rise of innovative technologies that provide reliable results as soon as a patient or user inputs their symptoms. Rather than making an appointment and visiting a hospital, now you can get accurate solutions online from the comfort of your own home. In this way, technology has made it possible for everyone to receive timely medical assistance whatever their needs may be. This system serves as a suggestion system to doctors and the health industry for the treatment of the patient, and any subsequent actions regarding the treatment are completely the responsibility of the doctors. Data sets from different websites about health have been gathered for the system that predicts diseases. The customer will be able to figure out how likely it is that using the signs and symptoms of a disease. Nowadays, the health industry is a big part of how people get better from their illnesses. This will help the health industry get the word out and help the user if they don't want to go to the hospital or any other clinics. This system tells doctors and the health industry how to treat a patient, and only the doctors can decide what to do next.

These study focus on disease like Diabetes, Heart Disease, Breast Cancer, Kidney Disease, Hepatitis.

Diabetes: Diabetes is a metabolic disorder that affects millions worldwide. It can cause high blood sugar levels and if left unchecked, result in serious health complications such as vision loss or heart disease. Diabetes falls into two main categories – Type 1, in which the immune system targets the cells that produces Insulin, and Type 2, in which insulin resistance leads to high blood sugar levels. Of these, it's estimated that 90-95% of cases are due to Type 2 diabetes - making it one of today's most significant global public health issues.

Heart Disease: Heart disease is a dangerous and pervasive health issue that affects the heart, blood vessels and even has an effect on our lives. Coronary artery disease being one of the most common types of such diseases usually builds up slowly as plaque formations block arteries supplying essential nutrients

to the heart muscles; in turn creating more severe issues with time if left untreated.

Hepatitis: Our livers are the hardworking organ responsible for keeping us healthy, but can be impacted by several forms of hepatitis. This dangerous condition is typically caused by viruses such as Hepatitis A, B and C or from excessive alcohol consumption, pollutants in our environment and certain drugs or medical disorders. It's important to know how to protect ourselves from this inflammation that affects our liver function - a vital element of overall wellbeing.

Breast Cancer: Breast cancer is a devastating disease that can affect anyone, regardless of age or gender. It starts when cells in the breast start growing and multiplying abnormally - quickly dividing out of control until they spread to other areas throughout the body via blood vessels and lymph nodes. If left untreated it can become metastatic with far-reaching consequences for those who suffer from it.

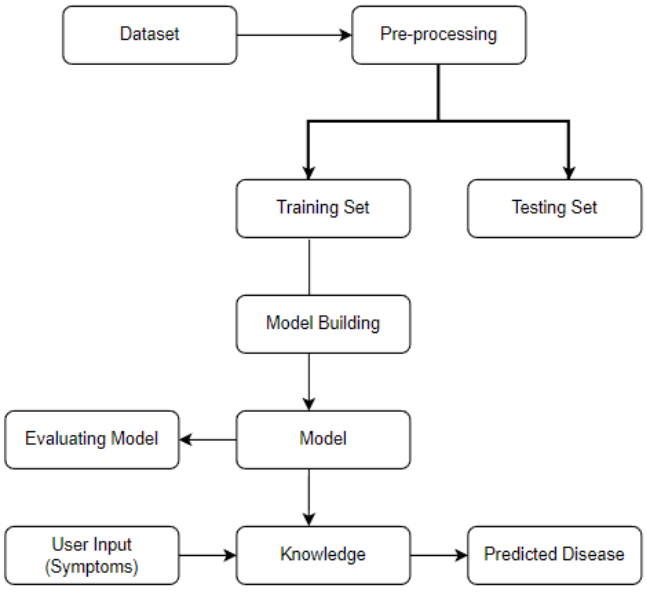
Chronic Kidney Disease: Chronic renal disease refers to a group of conditions that are harmful to the kidneys and diminish their ability to keep a healthy body by filtering waste from the blood. High blood pressure, anaemia (a low blood count), bone fragility, an unhealthy diet, and nerve damage are all potential complications. There is an increased risk of getting heart disease and blood vessel disorders in those who have kidney illness. These problems may become apparent slowly but surely over the course of time. Chronic renal disease may frequently have its progression halted if it is diagnosed and treated in a timely manner.

2. MATERIALS AND METHODS

2.1 . System Architecture

The purpose of this system is to identify cases of potentially fatal diseases and provide early analytic insights that can help physicians anticipate patterns in disease trends. Using both data mining and machine learning algorithms, the platform is able to quickly analyze patient information to predict possible outcomes - giving doctors an opportunity for timely intervention before it's too late.

A block diagram of multiple disease prediction is shown in Figure 1.



2.2. Data Preprocessing

Healthcare data may have impurities and missing values, decreasing its usefulness. As a result, data preparation is necessary to improve the efficacy and quality of the outcomes. This phase is required to achieve the desired result when employing machine learning approaches. The dataset requires preprocessing.

Missing values in the dataset can be estimated or removed. The most common way to deal with missing data is to fill it with the corresponding feature's mean, median, or mode. Because object values cannot be used for parsing, numerical data in object must be converted to float64. When dealing with categorical attributes, values that are null are replaced with the value that appears in the attribute column the most frequently. Label encoding is a method that associates each distinct value of an attribute with a numerical value. This method is used to convert category data into numeric attributes.

After replacing and encoding the data, it should be trained, validated, and tested. The process of training an algorithm to develop a model is called data training. A validation subset of the dataset is used to validate or improve various model fits. Use data tests to test model hypotheses.

3. METHODOLOGY

Data mining is one of the most essential and widely utilized technologies in the industry today for performing data analysis and gaining insight into the data. Approaches to data mining include Artificial Intelligence, Machine Learning, and statistical analysis, among other things. In this work, a machine learning technique is utilized to predict disease. Machine learning is a collection of tools and techniques that computers may use to turn unstructured data into useful information.

There are two types of machine learning algorithms in use right now. In supervised learning, problems with classification and regression are ubiquitous. It's widely used in predictive analytics since it creates a model from data that includes outcomes or reactions. The model is trained using labelled data. Unsupervised learning is used when the outcome or answers are uncertain, and the model is trained using unlabeled data. Two of the most common uses of this type of learning are pattern recognition and descriptive modelling.

This study discusses the use of different machine learning methods to predict disease. These algorithms include Naïve Bayes, K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Decision Tree, Logistic Regression, Random Forest, AdaBoost, and XGBoost is another technique.

3.1 Logistic Regression

Logistic regression is a widely used technique used for building machine learning models by analyzing data to detect relationships between independent and dependent factors. Data can be organized into nominal, ordinal or interval values which are then analyzed using the sigmoid function - an advanced cost function in contrast to linear regressions more basic approach. Logistic Regression provides valuable insights on both classification and regression type problems giving it broad capability across many different use cases.

3.2 KNN

KNN is an effective nonparametric machine learning technique that classifies objects based on their proximity to neighboring points. It takes into consideration an adjustable number of marked data-points (known as the K parameter) when deciding which category or value an unknown object should belong to, making it useful for both classification and regressions tasks alike. Despite its "delayed" effect - i.e., not immediately utilizing training set knowledge in order to determine what type something belongs with– KNN offers a versatile option within supervised artificial intelligence modelling techniques. The K-Nearest Neighbor algorithm is often referred to as the "lazy learner" due to its unique approach of waiting until it's time for classification before putting in any real work. Instead, the dataset is stored and only used when needed.

3.3 Random Forest

Random Forest is a commonly used supervised learning technique, combining the power of ensemble learning and decision trees to create a powerful model that accurately predicts outcomes while effectively avoiding overfitting. By utilizing multiple classifiers on subsets of data simultaneously, The Random Forests algorithm has established itself as one of the most reliable technique available for solving classification and regression issues within machine learning systems.

3.4 SVM

Support vector machines is a commonly used technique for tackling complex supervised learning tasks. By transforming the input data into an n-dimensional space using specialized kernels, SVM algorithms can accurately determine decision boundaries and classify previously unseen points without issue. By transforming the input data into an n-dimensional space using specialized kernels, SVM

algorithms can accurately determine decision boundaries and classify previously unseen points without issue. These kernel functions come in three varieties: linear, polynomial or radial basis — allowing them to find the best boundary even when problems cannot be broken up easily.

3.5 Decision Tree

Decision trees offer an efficient and effective way of mapping data to insights. By using a tree-like structure, they can quickly sort classification problems into neatly organized categories based on the records' attributes. As such, these models are commonly utilized by many industries in order to uncover new perspectives from their datasets. Choice nodes, which can have multiple branches, are used for decision making. The most difficult problem in developing a decision tree is determining the optimal attributes of the root node and sub-nodes. To overcome such challenges, we use information gain and the Gini index approach.

3.5 Naive Bayes

Naive Bayes is a powerful supervised learning technique, helping solve complex classification problems by harnessing the power of Bayes Theorem. Through probability-based predictions, it can tackle challenging tasks with greater accuracy and efficiency than simple models alone. This classifier determines where an example will fall into one of several categories. When it comes to data preparation, the Naive Bayes technique is known for its great precision and quickness. It's typically applied to large datasets. After steps of execution, characterization, estimation, and forecast, the Naive Bayes Algorithm is a probabilistic calculation that is sequential. Its advantage is that it does not necessitate a huge amount of training data.

3.6 Ada-boost

Adaptive Boosting (AdaBoost) is an innovative approach to creating ensembles that can consistently achieve higher accuracy than any one low-performing model. It works by setting classifier weights and training the data samples at each step, ultimately resulting in a single-level decision tree—or Decision Stump. With this method of constructing iterative models which are trained interactively on different weighted samples, AdaBoost provides consistent high-performance predictions for detection of anomalous observations.

AdaBoost should fulfill two requirements: Interactively training the classifier on various weighted training samples is recommended, with each iteration aiming to fit the samples as closely as possible while reducing training error.

3.7 XG-Boost

XGBoost is a powerful and reliable open-source approach for predictive analytics that combines the power of hundreds or thousands of weak models into one strong model. An efficient implementation of gradient-boosted trees, XGBoost utilizes a combination of convex loss function and regularization terms such as L1 and L2 to minimize error - creating highly accurate predictions with minimal effort.

The iterative training process creates successive regression trees which are then merged together for a consolidated response; making it possible for organizations in any field to unlock potential insights from their data sets more quickly than ever before

4. LITERATURE SURVEY

4.1. Diagnosing the Stage of Hepatitis C Using Machine Learning [1]

Hepatitis C is quite common around the world. The number of new cases of hepatitis C that are reported each year around the globe is between 3 and 4 million. globe. With the use of reliable and up-to-date disease prediction, individuals can obtain information on the stage of their Hepatitis C infection by using a variety of non-invasive blood biochemical indicators in conjunction with patient clinical data. Machine Learning techniques have proven to be a successful alternative method for identifying the stage of this chronic liver disease and preventing side effects of a biopsy. This study presents an exciting breakthrough in the field of healthcare diagnosis: an intelligent hepatitis C stage

system (IHSDS) powered by machine learning. By carefully selecting 19 features out of 29 from the UCI library, this IHSDS was able to achieve remarkable results when compared against earlier models; its precision value soared to 98.89% during training and 94.44% throughout validation! These truly promising figures highlight the potential for such powerful machine-learning tools as diagnostic aids within medical practice - potentially allowing us to improve on traditional methods with greater accuracy than ever before.

4.2. Multiple Disease Prediction System [2]

In the modern world, artificial intelligence and machine learning are extremely important. We can find them everywhere, from autonomous vehicles to the medical industry. Medical professionals are now leveraging the power of machine learning to create a multi-disease prediction system that takes immense amounts of patient data and processes it with unprecedented accuracy, bringing us ever closer to ensuring our patients stay healthy. Instead of relying on outdated methods capable only of predicting one disease at a time - often not correctly - we can rest assured knowing this new system offers reliable predictions every single time. Heart, liver, and diabetes are the three ailments we are now thinking about, but there may be many more added in the future. The system will output if the user has the disease or not when the user enters several disease-related parameters. Numerous people could benefit from this research because it allows for condition monitoring and medication administration. This system analysis is important because it takes into account all of the components that contribute to the onset of the diseases that are being studied, which makes it possible to identify them more efficiently and accurately.

4.3. Multiple disease prediction using Machine learning algorithms [3]

Interdisciplinary data mining is revolutionizing the healthcare sector by providing a fast and efficient way to evaluate medical treatments. By studying past database statistics, this field of research has allowed us to gain insights into many diseases, such as heart issues related to diabetes - an illness that disproportionately affects those living with it. However, current classification techniques are still unable to accurately predict cardiac problems for diabetics due its complexity in nature.

Cleveland data collections are used by the framework. The data is pre-processed before being put into machine learning algorithms such as SVM, KNN, and Decision Tree. Classified data flows into these algorithms. The prediction then uses the classification data as training data. The above approach has a Navies-Bayes accuracy of 0.77, an SVM accuracy of 0.87, and a decision tree accuracy of 0.90. Decision trees are generally considered the most reliable method for these datasets.

4.4. Prediction of Hepatitis Disease Using Machine Learning Technique [4]

The human liver plays an important role in the body. One of the illnesses that seriously impairs liver function is Hepatitis, a serious illness that causes inflammatory response in the liver. Hepatitis is a deadly disease that affects people all over the world. Vital body functions can be jeopardized if appropriate measures are not taken promptly. The disease is cured if it is detected early and correctly diagnosed. The datasets needed for this work are drawn from the UCI repository and selected with various clinical scenarios in mind. There are 155 instances in this dataset and 20 attributes. In this data set, one of the same attributes determines how long patients with hepatitis survive.

On the basis of accuracy, a comparative study of various techniques is carried out. We used SVM, KNN, and neural networks in this work. We achieved an accuracy of 89.58% using SVM, an accuracy of 85.29% using KNN, and an accuracy of 96.15 percentage using ANN, resulting in a good prediction for disease diagnosis.

4.5. Kidney Disease Diagnosis Based on Machine Learning [5]

Kidney disease is difficult to recognize. In addition, not only there is a lot of medical data on kidney disease, but it also has a complex structure. Methods for extracting and processing large amounts of medical data must become more efficient and comprehensive. In addition, machine learning technology can help medical professionals improve the accuracy of disease diagnosis and predict disease risk in patients.

The prediction dataset was obtained from the UCI repository containing 400 samples from kidney disease patients in India. Patients are classified according to the presence or absence of kidney disease. This dataset contains 250 patient records and 150 patients. Age, gender, total bilirubin, direct bilirubin, total protein and albumin are included in the 25 indicators. These datasets were then used to compare predictive outcomes for disease diagnosis using logistic regression and BP neural network models. Logistic regression is 89% and BP is 96%. As a result, the BP neural network outperformed logistic regression in diagnosing renal disease.

4.6. Kidney Disease Diagnosis using Classification Algorithm [6]

Recently, kidney disease has become a big problem. Diabetes, high blood pressure, leading an unhealthy lifestyle, eating junk food, being dehydrated, having arterial disease, and other factors are the primary contributors to kidney disease. Blood tests are used to diagnose kidney disease in its early stages. Routine test results make it difficult for doctors to make accurate diagnostic decisions. Machine learning algorithms are highly effective in predicting and analysing patient data during early stages of disease.

The mean imputation method handles missing values in the dataset and uses the Label Encoder method to convert the nominal dataset to a numeric dataset. Naive Bayes classifiers, K nearest neighbours(K-NN), bagging, decision trees, and random forest classifiers are applied to the training data set and test results are obtained and recorded. The results demonstrate that the decision tree algorithm enhanced kidney disease diagnostic accuracy by 98%.

4.7. Identification and Prediction of Chronic Diseases Using Machine Learning Approach [7]

Making an accurate diagnosis is a difficult process, and it can be even more challenging to predict when someone might suffer from certain chronic diseases. In this paper, we propose using cutting-edge machine learning techniques in order to accurately identify those who are at risk of developing such illnesses. To ensure that the categorization system works effectively, data mining will play a vital role. We will leverage CNNs for automatic feature extraction as well as disease prediction while KNN algorithms provide calculations on how close or far away our results are from exact matches in existing datasets. The data set was created by compiling illness symptoms, a person's lifestyle, and information related to doctor consultations. These factors are all taken into consideration throughout this generic prediction and diagnosis. Finally, this paper compares the suggested approach with a number of techniques, including Logistic regression, decision trees, and Naive Bayesian.

4.8. Prediction of Breast Cancer, Comparative Review of Machine Learning Techniques, and Their Analysis [8]

Breast cancer-related tumors can occur in breast tissue. It is one of the leading causes of mortality for women worldwide and the most common kind of cancer in females. In order to predict breast cancer, this article compares and contrasts data mining, deep learning, and machine learning methodologies. Breast cancer diagnosis and prognosis have attracted the attention of numerous researchers. Machine Learning and Data Mining are powerful tools for predicting outcomes, such as the presence of breast cancer. This study aims to explore a variety of current approaches in order to identify the most effective method that can accurately process large datasets. We look at existing research on ML algorithms used for this purpose, so beginners will gain an understanding of how these systems work and be equipped with knowledge relevant to deep learning models too. Our aim is not only compare/contrast existing strategies but also identify those methods best suited for working through huge datasets with extraordinary accuracy rates.

4.9. A Machine Learning Model for Early Prediction of Multiple Diseases to Cure Lives [9]

It is preferable to predict illness based on symptoms as opposed to consulting a physician. This lessens the pressure on the medical personnel and OPD wait times. We propose a system that is simple, inexpensive, and timesaving. Our proposed system is revolutionizing the connection between doctors and patients, enabling them to better collaborate in attaining their goals. Utilizing advanced algorithms such as Decision Trees, Naive Bayes, Random Forests and Artificial Neural Networks (ANN), it can accurately predict diseases based on a mere five user-specified symptoms - transforming traditional medical diagnoses for improved efficiency.

Our system examines and merges more datasets from different populations to produce more effective results. The system thus refines and improves the results. The accuracy for the above model is Decision Tree: 95.12%, Random Forest: 95.11%, Navy Bayes: 95.21%, ANN: 95.12%. In addition, our system's database is an incredibly useful resource for healthcare professionals, enabling them to store vital patient information and quickly recall existing diseases. It's a powerful tool in providing the most up-to-date treatment possible. It is a system that helps to improve health and make health management easier.

4.10. Multi Disease Prediction System [10]

A multi-illness prediction system built on predictive modelling uses the patient's symptoms as input to forecast their condition. The system calculates the possibility that the disease may manifest itself based on the patient's symptoms as input. The Random Forest Classifier is used to make predictions, and we also use Deep Learning Models (CNN) to make predictions for Malaria and Pneumonia. This method is more accurate, and we develop a web allocation system for prediction systems. This paper explored the effectiveness of Naive Bayes and five other classifiers on real-world medical tasks, using 15 datasets culled from the UCI Machine Learning Library. Logistic Regression, Decision Tree (DT), Neural Network (NN) and a simple rule-based method were all compared side by side to observe their performance in comparison with that of Naive Bayes. The experimental results showed that Naive Bayes' predicted accuracy results are superior to those of other methods in 8 of the 15 data sets. It was found that treating chronic illness on a worldwide scale is inefficient in terms of both time and money. Thus, the authors conducted this research to make predictions about the likelihood of sickness.

5. Evaluation Metrices

The three most crucial evaluation criteria are precision, recall, and accuracy. The four cardinalities of the confusion matrix can be used to define all of these measurements: The four categories are True positives (TP), False Positives (FP), False negatives (FN) and True negatives (TN). The ratio of true predictions for a class is used to assess **precision**, while the ratio of TP and (TP+FN) is used for calculating **recall**. The percentage of cases correctly categorized is known as **accuracy**. The sensitivity and specificity of the data were also considered. The ratio of correctly identified positive to correctly classified negative forecasts determines **sensitivity**, while the ratio of correctly classified negative to correctly classified positive forecasts determines **specificity**.

$$Precision = \frac{TP}{(TP + FP)}$$

$$Recall = \frac{TP}{(TP + FN)}$$

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)}$$

$$Sensitivity = \frac{TP}{(TP + FN)}$$

$$Specificity = \frac{TN}{(FP + TN)}$$

6. CONCLUSION

This study examined the potential of using a Machine Learning algorithm to accurately predict several diseases and forecast future illnesses based on patient-provided symptoms. It was believed that this cutting-edge methodology could reduce rates of chronic conditions by detecting them earlier, as well as bring down costs related to diagnosis and treatment for those affected. The findings suggest an innovative new system with no user threshold – meaning anyone can benefit from its use. By raising awareness about early-stage signs or signals of disease, we are one step closer towards helping people detect potentially life-changing medical problems in their earliest stages before it's too late.

7. REFERENCES

[1] Butt MB, Alfayad M, Saqib S, Khan MA, Ahmad M, Khan MA, Elmitwally NS. Diagnosing the stage of hepatitis C using machine learning. Journal of Healthcare Engineering. 2021 Dec 10;2021.

[2] Singh A, Yadav A, Shah S, Nagpure R. Multiple Disease Prediction System. International Research Journal of Engineering and Technology 2022.

[3] Arumugam K, Naved M, Shinde PP, Leiva-Chauca O, Huaman-Osorio A, Gonzales-Yanac T. Multiple disease prediction using Machine learning algorithms. Materials Today: Proceedings. 2021 Aug 2.

[4] Arumugam K, Naved M, Shinde PP, Leiva-Chauca O, Huaman-Osorio A, Gonzales-Yanac T. Multiple disease prediction using Machine learning algorithms. Materials Today: Proceedings. 2021 Aug 2.

[5] Wang B. Kidney Disease Diagnosis Based on Machine Learning. In2020 International Conference on Computing and Data Science (CDS) 2020 Aug 1 (pp. 307-310). IEEE.

[6] Pushpalatha S, Stella A. Kidney Disease Diagnosis using Classification Algorithm. In2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC) 2021 Nov 11 (pp. 1285-1288). IEEE.

[7] Alanazi R. Identification and prediction of chronic diseases using machine learning approach. Journal of Healthcare Engineering. 2022 Feb 25;2022.

[8] Fatima N, Liu L, Hong S, Ahmed H. Prediction of breast cancer, comparative review of machine learning techniques, and their analysis. IEEE Access. 2020 Aug 14;8:150360-76.

[9] Kumar A, Pathak MA. A machine learning model for early prediction of multiple diseases to cure lives. Turkish Journal of Computer and Mathematics Education (TURCOMAT). 2021 Jun 5;12(6):4013-23.

[10] Mandem D, Prajna B. Multi Disease Prediction System. International Journal of Innovative Research in Technology. 2021.

[11] Dutta S, Bandyopadhyay SK. A voting ensemble approach for hepatitis disease detection. medRxiv. 2020 Jan 1.

[12] Kumar N, Narayan Das N, Gupta D, Gupta K, Bindra J. Efficient automated disease diagnosis using machine learning models. Journal of Healthcare Engineering. 2021 May 4;2021.

[13] Kokol P, Jurman J, Bogovič T, Završnik T, Završnik J, Blažun Vošner H. Supporting Real World Decision Making in Coronary Diseases Using Machine Learning. INQUIRY: The Journal of Health Care Organization, Provision, and Financing. 2021 May;58:0046958021997338.

[14] Chavda K. Design and Development of Online Portal for Diagnosis of Multiple Diseases using AI. Int J Res Appl Sci Eng Technol. 2019 Mar 31;7(3):555–60.

[15] Sruthi G, Ram CL, Sai MK, Singh BP, Majhotra N, Sharma N. Cancer Prediction using Machine Learning. In: Proceedings of 2nd International Conference on Innovative Practices in Technology and Management, ICIPTM 2022. Institute of Electrical and Electronics Engineers Inc.; 2022. p. 217–21.

[16] Shaji S, Ajina SR, Priya R VS. Chronic Kidney Disease Prediction using Machine Learning.

[17] Kunjir A, Sawant H, Shaikh NF. Data mining and visualization for prediction of multiple diseases in healthcare. In2017 International Conference on Big Data Analytics and Computational

Intelligence (ICBDAC) 2017 Mar 23 (pp. 329-334). IEEE.

[18] Moreno-Ibarra MA, Villuendas-Rey Y, Lytras MD, Yáñez-Márquez C, Salgado-Ramírez JC. Classification of diseases using machine learning algorithms: A comparative study. Mathematics. 2021 Jul 31;9(15):1817.

[19] Uddin S, Khan A, Hossain ME, Moni MA. Comparing different supervised machine learning algorithms for disease prediction. BMC medical informatics and decision making. 2019 Dec;19(1):1-6.

[20] Kate RJ, Perez RM, Mazumdar D, Pasupathy KS, Nilakantan V. Prediction and detection models for acute kidney injury in hospitalized older adults. BMC medical informatics and decision making. 2016 Dec;16(1):1-1.