IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

DISEASE PREDICTION AND MEDICATION ADVICE USING MACHINE LEARNING ALGORITHMS

Pooja Panapana[1], K. Sri Rakesh Reddy[2], J. Deepika[3], G. Rushivardhan Babu[4] and A. Drakshayani[5]

[1]*Information Technology, GMR institute of Technology, India*
Email: pooja.p@gmrit.edu.in
[2]*Information Technology, GMR institute of Technology, India*
Email: rakeshreddy.karri@gmail.com
[3]*Information Technology, GMR institute of Technology, India*
Email: deepikajinugu781@gmail.com
[4]*Information Technology, GMR institute of Technology, India*
Email: rushivardhan10@gmail.com
[5]*Information Technology, GMR institute of Technology, India*
Email: drakshayaniadduri@gmail.com

ABSTRACT :

People today deal with a variety of diseases as a due to their lifestyle choices and the surroundings Disease prediction is an essential component of treatment. So, it becomes crucial to make disease predictions early on. The hardest task is making an accurate diagnosis of a disease. So, Machine learning is crucial in predicting the disease in order to solve this issue. The system proposed in this project uses the patient's symptoms as input to predict the disease and then recommends the right medication. This system takes the symptoms of the user from which he or she suffers as an Input. We use Classification Algorithms for disease prediction and drug recommendation, such as Naive Bayes (NB) and Random Forest, with a variety of accuracy levels. Once the disease has been predicted by the system, then it is followed by recommending the appropriate medicine. In order to improve the capabilities of the current systems, this paper discusses about the creation of a system that serves the dual purpose of disease prediction and medication suggestion. To make it easier for users to engage with the symptoms, an interactive interface is designed as the front-end and the front-end is deployed by using Flask.

Keywords: Naive Bayes (NB), Random Forest, Disease Prediction, Drug Recommendation, Flask, Healthcare.

1. INTRODUCTION:

Health and pharmaceutical are going gradually important in today's rapidly changing society, as arising technology is being employed to battle virtually each known conditions. Due to the current state of the surroundings and their life choices, humans are now exposed to a wide range of conditions. According to estimations, more than 70% of Indians are exposed to common disorders including the flu, cold wave, cough, and viral infections for every two months. Since numerous people are ignorant that general health problems could be lateral goods of reality more dangerous. The absence of early discovery of numerous conditions, including cancer, diabetes, and others, is the primary reason for adding the cause of deaths on a global scale. Therefore, it is important to identify the illness as early as possible to maintain a strategic distance from any unfavourable losses. In order to identify the diseases early, Machine learning plays an important role. Machine Learning is the understanding of a computer system in which the Machine Learning model learns from data and experience. Training and Testing are the two iterations of the machine learning algorithm. ML model allows us to build models to get quickly cleaned, processed data, and deliver results faster. The suggested system predicts the disease from the symptoms data set and provides a general medication advice based on the patient's symptoms using machine learning methods such as Gaussian Naive bayes Algorithm and Random Forest Classifier.

2. LITERATURE SURVEY:

To create this project, we looked at ten publications from various external sources. We examined the relationship between various algorithms' performance in various disease prediction scenarios.

This paper looks at the use of Machine Learning to develop a disease prediction and doctor recommendation system. Different classification algorithms like Logistic Regression, Random Forest Classifier, KNN and Naïve Bayes are used to predict a person's disease based on their symptoms and then recommends which type of doctor to consult. This system is used by end-users. An interactive interface is built as front-end and is connected to the Server. This system might have a significant impact on how doctors treat patients in the future. But, due to the complexity and variety of diseases, there may be possible accuracy issues as well as bias in the data used to train the algorithm.[1]

In this study, we discuss the application of machine learning to predict diseases from patient symptoms. It determines the probability of what disease could be present by using the supervised machine learning algorithm like Naive Bayes classifier. Accurate analysis can help in early identification and improve patient care along with the development of biomedical and healthcare data. It also shows how the linear regression and decision tree algorithms can be used to predict specific diseases like Diabetes, Malaria, Jaundice Dengue or Tuberculosis. The advantages of using Machine Learning model includes the ability to accurately analyse medical data, leading to early detection and better patient care. Additionally, it can be used to predict specific

diseases with a high degree of accuracy. However, this approach needs a lot of data for the algorithms to function well, and bias might result if insufficient varied datasets are used.[2]

The standard method of diagnosis involves a patient seeing a doctor, going through several tests, and then coming to a decision. It takes a long time to do this task. This project suggests an automated disease prediction system to reduce the time needed for the initial disease prediction process, which depends on user input. It is designed in a way such that when the user is introduced to the chatbot system, they are given the choice of receiving an estimation or prediction of their disease based on the data they have provided to the chatbot. Data is gathered from Columbia University to help with disease prediction, and Kaggle provides a source for the diseases' related symptoms. There are several machine learning algorithms used, such Naive Bayes, Random Forest Classifier, K-Nearest Neighbours, and Support Vector Machine classifier. However, when fewer symptoms are submitted, the accuracy will be reduced. [3]

This research paper proposes using data mining methods like K-Nearest Neighbour (KNN) and Convolutional Neural Network (CNN) machine learning to examine large amounts of medical data and find patterns in order to accurately predict diseases. The accuracy of using CNN for general disease prediction is 84.5%, which is better than using KNN, as well as taking less time and memory usage than KNN. After general disease prediction, the system can indicate how likely someone is to suffer from a certain disease. The advantages of using data mining methods like KNN and CNN for general disease prediction are that they can accurately predict diseases, with the accuracy of using CNN being 84.5%. Additionally, these methods take less time and memory usage than KNN. The disadvantage is that it requires a large amount of medical data to be able to find patterns in order to make accurate predictions. The main limitation of using data mining methods like KNN and CNN for general disease prediction is that it requires a large amount of medical data to be able to find patterns in order to make accurate predictions. [4]

The study examines the use of machine learning (ML) to computer-aided diagnosis (CAD), a field of medical research. It talks over how Machine Learning can be used to assure impartiality in decision-making processes and to increase the accuracy and reliability of disease identification. The study also looks at several machine learning (ML) algorithms and methods for identifying conditions like diabetes and heart disease and drawing conclusions based on. This approach may also be used to analyse huge datasets from several medical sources and forecast illnesses. However, the downside of using ML in CAD is the possibility of errors based on by inaccurate data or algorithmic assumptions. The amount of data that can be used and evaluated may also be constrained, and the dataset may contain biases.[5]

In this study, they proposed an interactive system for predicting thyroid disease that makes use of machine learning methods. For the prevention of thyroid, several machine learning methodologies and diagnoses are also proposed. The estimated probability of a patient getting thyroid illness is predicted using a variety of machine learning algorithms, such as support vector machine (SVM), K-NN, and decision trees. The entire data set describing the symptoms was collected from the UCI machine learning library.[6]

Effective treatment is greatly aided by an intelligent system that can predict diseases accurately and recommend appropriate medical facilities. This system analyses patient symptoms include heart rate, blood pressure, and blood sugar level information that is input into the system while it is running to predict the disease with the highest degree of accuracy. The system suggests nearby hospitals and clinics based on the prediction and the user's choices. Three algorithms—Naive bayes, Weka, and core NLP are compared in this work. The disease indicated by the Naive Bayes classifier stands out among these systems, with an estimated accuracy of more than 80%.[7]

In recent years, heart diseases have become one of the main causes of death and the deadliest disease globally. The study examines how the Support vector machine algorithm may be used to predict heart disease by taking into consideration many factors, including age, gender, age, cholesterol, and heart rate. The heart disease data set was collected from the UCI machine learning repository. It forecasts a disease risk level between 0 and 1, with 0 indicating no risk and 1 suggesting a high risk, indicating the presence of danger.[8]

For the purpose of preventing and treating illness, accurate and prompt examination of any health-related issue is important. This system creates a medical diagnostic system based on machine learning (ML) algorithms for illness prediction, which can help in more accurate diagnosis than the traditional way. Several ML models are used, which includes Fine, Medium, and Coarse Decision trees, Gaussian Naive Bayes, Kernel Naive Bayes, Fine, Medium, and Coarse KNN, Weighted KNN, Subspace KNN, and RUS Boosted trees. They reported around 230 disorders in the dataset, with over 1000 distinct symptoms in total. These machine learning algorithms were evaluated with the symptoms, age, and gender of a person as input. They used the Weighted KNN method and obtained the maximum accuracy of 93.5% for disease prediction.[9]

This research paper uses data from a dataset of 303 heart disease patients from the Cleveland database of UCI repository, as well as supervised learning algorithms such as Naïve Bayes, decision tree, K-nearest neighbour and random forest algorithm. The goal is to estimate the likelihood of a person developing heart disease by testing only 14 attributes. The results show that the highest accuracy score was achieved using the K-nearest

neighbour algorithm. The advantages of using data mining and machine learning techniques to analyse medical data is that it can help healthcare professionals to predict heart disease more accurately. The disadvantage is that only 14 attributes are being tested, which may not be enough for a reliable prediction.[10]

This paper reviews existing literature on the use of intelligent computing and big data analytics in healthcare, specifically focusing on drug recommendation systems. It looks at how machine learning can be used to develop personalized drug recommendations for each patient's health condition based on ratings and reviews from other users. Additionally, it examines the potential benefits of using sentiment analysis and visualization techniques to improve accuracy. The advantages of this paper include the potential to improve accuracy and efficiency in healthcare decisions by providing personalized drug recommendations for each patient's health condition. Additionally, it can help reduce errors in medication that could be a threat to patients' lives. The disadvantages of this paper include the difficulty of obtaining accurate data from different websites on the Internet, as well as potential bias or incorrect information due to sentiment analysis and visualization techniques used.[11]

The paper discusses a study on using machine learning algorithms to predict diseases in healthcare communities, with a focus on treatment task queue management. Machine learning is a field of software engineering that allows computers to learn and improve performance on a particular task without being programmed explicitly. In this paper, the system uses structured and unstructured data and algorithms such as decision tree and KNN to predict various disorders and waiting times for treatment tasks for patients. The accuracy of the proposed algorithm is 94.8%, and it outperforms several standard estimate algorithms. Patient queue management is challenging and complicated, as patients may require different phases, and each task may have varying time requirements. The paper aims to develop a patient treatment time consumption model to help hospitals schedule each treatment task queue and avoid overcrowded and ineffective queues. [12]

This article discusses a research project that uses machine learning algorithms to predict diseases based on a patient's symptoms and recommend appropriate medicine. The project uses Decision Tree, Random Forest, and Naïve Bayes algorithms to achieve its objective. The article also provides information on the system architecture, the dataset used, and the algorithms used. The proposed system aims to improve healthcare by accurately and early detection of diseases, reducing the workload on medical staff and improving patient care. The system is found to have an accuracy of 90% or above based on the confusion matrix.[13]

This paper surveys the literature of existing deep machine learning methods for predicting and recommending treatment for multidisciplinary diabetes disease. It compares these methods to its own proposed system, which uses optimal machine learning models and data fusion techniques on healthcare datasets. The advantages of this paper are that its proposed system achieved 99.6% accuracy, which is higher compared to the existing deep machine learning methods. Additionally, it can be used in automated diagnostic and recommendation systems for diabetic patients to improve disease diagnosis performance. The disadvantages of this paper are that it does not address other potential treatments or therapies for diabetes beyond what is discussed in the literature survey and its proposed system may not be applicable to all types of diabetes cases due to differences in individual patient conditions.[14]

The paper discusses a computational system for disease diagnosis and prescription generation using decision tree classification and affinity analysis. The system takes user-entered symptoms, uses association rules to identify relevant symptoms, predicts the disease, and generates a prescription based on user attributes and the identified disease. The system can predict multiple diseases if the probabilistic value of classes is similar and encourages users to seek medical attention. Future work includes expanding the dataset and developing a system for verification by doctors to build user trust.[15]

3. METHODOLOGY:

3.1 Random Forest: -

A Random Forest Method is a supervised machine learning algorithm that is widely used in Machine Learning for classification and regression issues. The more trees in a Random Forest Algorithm, the more accurate and problem-solving capacity it has. Random Forest is a classifier that uses the average of many decision trees on different subsets of a given dataset to increase its predicting accuracy. It is built on the idea of ensemble learning, which is the process of integrating numerous classifiers to solve a complex problem and improve the model's performance. Ensemble means the combination of multiple models. Ensemble mainly uses two types of methods: Bagging(parallel), Boosting(sequential). Random forest algorithms have three major hyperparameters that must be set before training begins. These parameters include node size, number of trees, and number of characteristics sampled. The random forest classifier can be used to solve regression or classification problems.

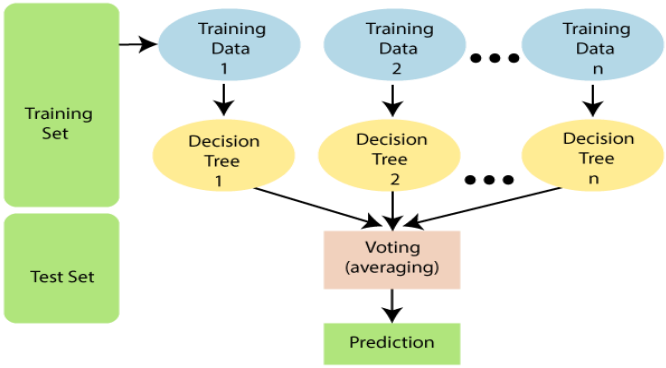

Fig3.5: Working of Random Forest

Working of Random Forest Algorithm

Step 1: Select random samples from a given data or training set.

Step 2: This algorithm will construct a decision tree for every training data.

Step 3: Voting will take place by averaging the decision tree.

Step 4: Finally, select the most voted prediction result as the final prediction result.

Here , From the disease dataset it will create a Bootstrap dataset i,e it will randomly picks up some symptom values from dataset and the values can be repeated. The dataset is divided into subsets and given to each decision tree. Each decision tree produces a prediction result . Instead of depending on a single decision tree, the random forest considers the predictions from each tree and predicts the final output based on the majority vote of predictions.

When compared to other algorithms, it requires less training time. It predicts output with excellent accuracy, and it works efficiently even with huge datasets. It can also retain accuracy when a huge portion of the data is missing. Disease patterns and disease risks may be recognised with the use of this technique.

3.2 Naive Bayes:

The Naive Bayes algorithm is a probabilistic machine learning algorithm that can be used for a wide range of classification tasks. It is based on Bayes Theorem. The word "Naive" is used since the algorithm incorporates features in its model that are independent from each other. Any changes in the value of one feature have no effect on the value of any other feature of the algorithm i,e it is assumed that the value of one feature is independent of the value of any other feature. In practise, naive bayes classifiers are often found to be quite efficient, especially when the number of dimensions of the feature set is large.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Fig 3.6: Bayes Theorem

To calculate the posterior probability P(B|A), first create a Frequency Table for each feature against the target. The frequency tables are then shaped into Likelihood Tables. Finally, calculate the posterior probability for each class using the Nave Bayesian equation. The class with the highest posterior probability is the outcome of the prediction

3.3 Gaussian Naive Bayes (GNB):

Gaussian Naive Bayes is a probabilistic classification algorithm that uses the Bayes theorem with strong independence assumptions. It is a variant of Naive Bayes that supports continuous data and corresponds to the Gaussian normal distribution. Each parameter (also known as a feature or a predictor) is assumed to have an independent ability to predict the output variable by Gaussian Naive Bayes. It is a fast and flexible model that produces highly reliable results on large data sets. There is no need to spend a lot of time towards training. It also improves grading performance by removing unimportant specifications.

The final prediction is the combination of the predictions for all parameters that returns the probability of the dependent variable would be placed in each group. The final classification is given to the group with the highest probability.

$$P(x_i \mid y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left(-\frac{(x_i - \mu_y)^2}{2\sigma_y^2}\right)$$

Fig 3.7: Formula for Gaussian Naive Bayes

The Gaussian probability density function can be used to create predictions by replacing the parameters with the new variable input value, and the Gaussian function will give an estimate for the probability of the new input value.

3.4 Logistic Regression:

Logistic Regression is a statistical technique used to predict a binary outcome based on a set of independent variables or predictors. It works by using a mathematical function called the logistic function or sigmoid function to estimate the probability of the binary outcome. The logistic function takes the values of the independent variables as input and maps them to a value between 0 and 1, which represents the probability of the binary outcome. The logistic

regression model uses a training dataset with known outcomes to estimate the values of coefficients that maximize the likelihood of the observed data.

The logistic regression model applies this function to a linear combination of the independent variables (also called features or predictors) to predict the probability of the binary outcome.

The equation for the logistic function is as follows:

p = 1 / (1 + e^-z)

where p is the probability of the binary outcome,

e is the base of the natural logarithm

z is the linear combination of the independent variables:

$z = \beta 0 + \beta 1x1 + \beta 2x2 + ... + \beta nxn$

where β0, β1, β2, ..., βn are the coefficients or weights assigned to each independent variable

x1, x2, ..., xn are the values of those independent variables.

Once the coefficients are estimated, the logistic regression model can be used to predict the probability of the binary outcome for new observations. If the probability is greater than or equal to a chosen threshold, the observation is classified as a 1; otherwise, it is classified as a 0. It is a simple and interpretable method that can be applied to a wide range of applications, such as medical diagnosis, fraud detection, and customer churn prediction.

3.5 K Nearest Neighbour (KNN):

K Nearest Neighbour (KNN) is a simple and popular machine learning algorithm used for classification and regression tasks. The main idea behind KNN is to find the k closest points (neighbours) in the training data to a new data point and use their labels to predict the label of the new point.

Here are the simplified steps involved in the KNN algorithm:

1. Collect the training data with known labels or values.

2. Choose a value for k, the number of nearest neighbours to consider.

3. Calculate the distance between the new data point and all the points in the training data.

4. Select the k nearest data points to the new data point.

5. For classification, predict the label of the new data point based on the majority vote of the k nearest neighbours' labels. For regression, predict the value of the new data point based on the average of the k nearest neighbours' values.

6. Evaluate the performance of the model by testing it on a separate set of data with known labels or values.

KNN is a non-parametric algorithm, which means that it does not make any assumptions about the underlying data distribution. It is also a lazy learning algorithm, which means that it does not build a model during training but instead stores the entire training data set for use during prediction. KNN is simple to implement and can work well on small datasets with a low number of features. However, it can be computationally expensive for large datasets with many features.

3.6 Support Vector Machine (SVM):

Support Vector Machine (SVM) is a supervised machine learning algorithm used for classification and regression tasks. The basic idea behind SVM is to find the best possible line or hyperplane that separates the data points into different classes. The line or hyperplane is chosen in such a way that it maximizes the margin between the two classes of data. The margin is the distance between the line or hyperplane and the closest data points of each class. The data points closest to the line or hyperplane are called support vectors.

SVM works by transforming the data into a higher-dimensional space using a kernel function. This transformation makes the data more separable in that space. The choice of kernel function depends on the data and the classification problem. SVM then finds the hyperplane that separates the classes with the maximum margin by solving an optimization problem.

Once the hyperplane is found, SVM can classify new data points by determining which side of the hyperplane they fall on. SVM is a powerful algorithm that can work well in a variety of classification tasks, but it can be sensitive to the choice of kernel function and parameters, and it may not work well with very large datasets.

Table 1: COMPARISION OF DIFFERENT METHODS

Title of the project	Algorithms and Techniques	Advantages	Disadvantages
[1] Disease Prediction and Doctor Recommendation System using Machine Learning Approaches	Logistic Regression, Random Forest Classifier, K-Nearest Neighbour (KNN), Naïve Byes.	It provides a way to accurately predict diseases by looking at the symptoms of patients, and then recommends which type of doctor to consult.	It include that machine learning technology can be expensive or difficult to implement in some cases due its complexity.
[2] The prediction of disease using machine learning	Decision Tree, Random Forest, Naïve Bayes, Support Vector Machine (SVM).	It provides a way to accurately predict diseases using Machine Learning, which can help doctors make better decisions about patient care.	The cost associated with implementing such a system, as well as potential errors in predictions due to incomplete or incorrect data input from users.
[3] Medical Disease Prediction using Machine Learning Algorithms	Multinomial NAÏVE BAYES, Random Forest Classifier, K-Nearest Neighbour Classifier, Support Vector Classifier.	It proposes a way to use machine learning algorithms to predict diseases, which could be helpful for people who have difficulty accessing medical care or need an initial baseline idea of their condition.	It may include potential inaccuracies in the predictions due to lack of data or incorrect assumptions about certain conditions and symptoms.

[4] Designing Disease Prediction Model Using Machine Learning Approach	K-Nearest Neighbour (KNN), Convolutional neural network (CNN).	It also uses K-Nearest Neighbor and Convolutional neural network machine learning algorithms to make predictions, which require less time and memory resources.	Include the need for a dataset of disease symptoms in order to make accurate predictions as well as consideration for living habits and check-up information when making diagnoses.
[5] Symptoms Based Disease Prediction Using Machine Learning Techniques	Support Vector Machine (SVM), Naive Bayes.	It provides a comprehensive overview of the various machine learning algorithms used for disease detection and decision-making processes	It include potential bias or errors due to relying on automated algorithms, which could lead to incorrect diagnoses or treatments if not properly monitored.
[6] Interactive Thyroid Disease Prediction System Using Machine Learning Technique	Support Vector Machine (SVM), Decision Tree, K- Nearest Neighbour.	It proposes a system for predicting thyroid disease using machine learning techniques, which can be used to help diagnose and prevent the onset of the disease.	It include that machine learning algorithms may not always accurately predict risk levels due to potential bias in datasets or other factors.
[7] Disease Prediction and Doctor Recommendation System	Naive bayes	The recommendation system suggests specialists based on user's chosen filters such as location, fees, experience, and feedback reviews.	There is potential risk associated with relying solely on automated systems for making decisions regarding patient care without human input from healthcare professionals.
[8] Heart disease prediction and Doctor recommendation system using machine learning	Support Vector Machine (SVM).	It proposes a system for accurately predicting heart diseases, which could greatly benefit future medical therapy.	It include potential inaccuracies in predictions due to incomplete or incorrect data sets used for training models
[9] Disease prediction from various symptoms using machine learning	Gaussian Naive Bayes, KNN.	It provides a more accurate diagnosis than the traditional method, as well as providing an early diagnosis to ensure timely treatment.	There is a risk that some diseases may not be included in the dataset used by this system which could lead to inaccurate results.
[10] Heart Disease Prediction using Machine Learning Techniques	Naïve Bayes' Classifier, K-Nearest Neighbour (K-NN), Decision Tree, Random Forest.	This can help healthcare professionals make an early diagnosis and prompt management for those at risk	Only 14 attributes out of 76 were considered for testing, so there may be other important factors not considered when making predictions

			about heart disease risk.
[11] A Machine Learning based Drug Recommendation System for Health Care	Logistic Regression.	It provides a way for users to make more efficient and accurate health-related decisions by using machine learning and data mining techniques.	It include potential bias in recommendations based on user reviews, ratings, or condition systems.
[12] Disease Prediction by Machine Learning from Healthcare Communities	K Nearest Neighbour, (KNN) and Support Vector Machine (SVM).	It include potential bias in recommendations based on user reviews, ratings, or condition as well as difficulty implementing such a system across different healthcare systems.	Incomplete medical data which may lead to reduced accuracy in predictions and regional diseases which could weaken outbreak predictions.
[13] Disease prediction and Prescription suggestion using Machine learning	Naive Bayes, Random Forest, Decision Tree.	This could help reduce medication errors and provide better patient care.	This paper may be the cost associated with implementing such a system, as well as potential accuracy issues due to the complexity of medical diagnosis.
[14] A Smart Healthcare Recommendation System for Multidisciplinary Diabetes Patients with Data Fusion Based on Deep Ensemble Learning	Naive Bayes, Random Forest, Decision Tree, K Nearest Neighbour, (KNN) and Support Vector Machine (SVM).	It proposes a deep machine learning and data fusion approach to accurately predict and recommend treatment for diabetes, which is more accurate than existing deep machine learning methods.	It does not address other diseases or medical conditions beyond diabetes disease, so its application may be limited in scope.
[15] A Computational System for Disease Diagnosis and Prescription Generation	Decision Tree Classifier.	The system provides reliable and adequate diagnoses, especially in rural areas where healthcare resources are scarce.	This system may not be able to accurately diagnose complex diseases or conditions that require specialized knowledge or expertise from medical professionals.

4. CONCLUSION:

This disease prediction system's primary goal is to make disease predictions based on symptoms. This system takes the user's symptoms as input and generates a result in the form of a disease prediction and fit gives medication advice based on the disease. This paper proposed a method for identifying and predicting the presence of a disease in an individual using machine learning algorithms such as Naive Bayes, Random Forest, K Nearest Neighbour, Gaussian Naive Bayes, Logistic Regression and Support Vector Machine. We discovered that the Support Vector Machine is the most used, algorithm followed by the Random Forest. Support Vector Machine produces the best results because it is faster and offers highest accuracy of 99.63%. It is widely assumed that the suggested method can reduce illness risk by identifying them early and lower the cost of diagnosis and treatment. However, the choosing of symptoms has a significant impact on disease prediction accuracy.

In the future, we can further enhance the model by including Deep Learning Algorithms and using vast datasets directly obtained from hospitals. To make the project more user-friendly, we can implement it entirely within the Android application.

5. REFERENCES:

[1] “Disease Prediction and Doctor Recommendation System using Machine Learning Approaches”. International Journal for Research in Applied Science and Engineering Technology. 9. 10.22214/ijraset.2021.36234.

[2] Gomathy, C K. (2021). “The Prediction of Disease using Machine Learning”.

[3] Kunal Takke, Rameez Bhaijee, Avanish Singh;"Medical Disease Using Machine Learning Algorithms";2022 International Journal for Research in Applied Science & Engineering Technology (IJRASET), May 2022

[4] 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, doi: 10.1109/ICCMC.2019.8819782..

[5] P. Hamsagayathri and S. Vigneshwaran, "Symptoms Based Disease Prediction Using Machine Learning Techniques," 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), 2021, pp. 747-752, doi: 10.1109/ICICV50876.2021.9388603.

[6] A. Tyagi, R. Mehra and A. Saxena, "Interactive Thyroid Disease Prediction System Using Machine Learning Technique," 2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC), 2018, pp. 689-693, doi: 10.1109/PDGC.2018.8745910.

[7] Dhanashri Gujar, Rashmi Biyani, Tejaswini Bramhane, P Vaidya "Disease Prediction and Doctor Recommendation System," March-2018 International Research Journal of Engineering and Technology(IRJET) .

[8] Juned Hakim Shaikh, Abhishek Ganesh Takale ,"Heart Disease Prediction and Doctor Recommendation System using Machine Learning," Nov-2022 International Research Journal of Modernization in Engineering Technology and Science(IRJMETS).

[9] Keniya, Rinkal and Khakharia, Aman and Shah, Vruddhi and Gada, Vrushabh and Manjalkar, Ruchi and Thaker, Tirth and Warang, Mahesh and Mehendale, Ninad and Mehendale, Ninad, Disease Prediction from Various Symptoms Using Machine Learning (July 27, 2020).

[10] Shah, D., Patel, S., & Bharti, S. K. (2020). Heart Disease Prediction using Machine Learning Techniques. SN Computer Science, 1(6). doi:10.1007/s42979-020-00365-y\

[11] Mohapatra, Mahima; Nayak, Mamata; and Mahapatra, Saswati (2022) "A Machine Learning based Drug Recommendation System for Health Care," *Graduate Research in Engineering and Technology (GRET)*: Vol. 1: Iss. 6, Article 2. DOI: 10.47893/GRET.2022.1109

[12] Jadhav, Saiesh & Kasar, Rohan & Lade, Nagraj & Patil, Megha & Kolte, Shital. (2019). Disease Prediction by Machine Learning from Healthcare Communities. International Journal of Scientific Research in Science and Technology. 29-35. 10.32628/IJSRST19633.

[13] Ashutosh Singh,Himanshu Gupta,Mihir Kumar,Rishabh Agrawal,Tejaswini K.(2022).DISEASE PREDICTION AND PRESCRIPTION SUGGESTION USING MACHINE LEARNING.International Research Journal of Modernization in Engineering Technology and Science(IRJMETS).

[14] Baha Ihnaini and M. A. Khan and Tahir Abbas Khan and Sagheer Abbas and Mohammad Sh. Daoud and Munir Ahmad and Muhammad Adnan Khan. (2021).A Smart Healthcare Recommendation System for Multidisciplinary Diabetes Patients with Data Fusion Based on Deep Ensemble Learning.Computational Intelligence and Neuroscience Hindawi.10.1155/2021/4243700.

[15] Bakhrey*, N., Bakhtiani, R., Soni, J., & Kalbande, Dr. D. (2019). A Computational System for Disease Diagnosis and Prescription Generation. In International Journal of Innovative Technology and Exploring Engineering (Vol. 9, Issue 2, pp. 1906–1910). Blue Eyes Intelligence Engineering and Sciences Engineering and Sciences Publication - BEIESP. https://doi.org/10.35940/ijitee.b6262.129219