Predictive Modeling for Chronic Disease Management

Gauraangi Praakash*, Prof.Arnab Bhattacharya,Dr Pooja Khanna*****

*B.Tech CSE with Hons. In AIML Scholar Amity University Lucknow

**Professor, Dept. of Computer Science and Engineering, Indian Institute of Technology, Kanpur.

***Assistant Professor, ASET Amity University Lucknow

ABSTRACT

Chronic diseases present significant healthcare challenges worldwide, affecting mortality rates and straining healthcare budgets. Recent advancements in medical research have enabled the collection of health-related data, integrating structured and unstructured information to enhance accuracy. Machine learning techniques and intelligent healthcare solutions have emerged as valuable tools in mitigating risks, detecting diseases at an early stage, and effectively managing chronic conditions. To improve prediction accuracy, preprocessing methods have been developed to address missing data and select relevant features. Machine learning models leverage user-input symptoms and online activities to predict prevalent chronic diseases like diabetes, cardiovascular diseases, and cancer. Mobile applications empower patients to take charge of their health by facilitating self-management of their conditions. In this context, the proposed system utilizes K-Nearest Neighbors (KNN) and Convolutional Neural Networks (CNN) for the detection of chronic diseases. Overall, the integration of machine learning techniques and healthcare data shows great potential for disease forecasting and management. By harnessing the power of data analysis and predictive modeling, healthcare professionals can make more informed decisions, improve patient outcomes, and optimize resource allocation.

Keywords: chronic diseases, healthcare challenges, medical research, health-related data, machine learning techniques, intelligent healthcare solutions, prediction accuracy, preprocessing methods, feature selection, disease detection, mobile applications, self-management, K-Nearest Neighbors (KNN), Convolutional Neural Networks (CNN), disease forecasting, disease management.

1. INTRODUCTION

Chronic diseases, which last for a long time and need ongoing medical care and management, include cancer, heart disease, and diabetes. They are non-communicable diseases. Unfortunately, chronic diseases are a serious global health issue, and they are becoming more and more common. In order to effectively manage chronic diseases, it is crucial to identify potential risks and intervene in a timely manner. Major worldwide health challenges include chronic diseases including diabetes, cancer, and heart attacks. These conditions frequently call for long-term monitoring and therapy and can have a major negative influence on a person's quality of life. Early detection is crucial for effective management of these chronic conditions.

In order to create predictive models that can precisely detect and anticipate the start of chronic diseases, machine learning algorithms are being used more and more. To determine risk factors and forecast the start of diseases, these models make use of a variety of data sources, including clinical databases and medical records. Predictive models based on machine learning have been created for a number of chronic conditions, including type 2 diabetes, asthma, chronic kidney disease, arthritis, Alzheimer's disease, dementia, and HIV/AIDS, utilising data from medical records and clinical databases.

With the aid of these prediction models, risk factors that contribute to the development of certain illnesses can be identified, enabling early intervention and better management outcomes. In order to deliver individualized and focused care to people with chronic conditions, machine learning algorithms are also able to analyze vast volumes of complicated data from numerous sources, including biometric sensors and wearables. Predictive models for chronic disease detection and management have been made possible through the application of machine learning algorithms, which can dramatically increase. It is important to keep in mind that some researches have noted limits to the use of machine learning algorithms in forecasting the onset of chronic diseases since they may not be applicable or generalizable to all populations or may necessitate a considerable investment in computer power. However, as additional data sources become accessible, the application of machine learning algorithms in chronic disease prediction has showed promise and is still developing.

The main goal of the proposed system is to use machine learning to detect and predict chronic disease in an individual. The patient's information in the data set is both structured and unstructured, including unstructured information about symptoms, unstructured information about interactions with medical professionals addressing the condition, and unstructured information about everyday activities. The structured data excludes any patient-specific information, such as name and ID number. The missing values are discovered through pre-processing of these data. They are subsequently rebuilt to improve the model's quality and the model's predictability. Machine learning techniques like CNN and KNN are utilised for prediction.

2. RELATED WORK

This section outlines the associated research that was done in order to create the suggested model for chronic illness prediction.

Developing predictive models for chronic disease management requires a comprehensive understanding of the various data sources available. Machine learning techniques can be utilized in conjunction with the Common Data Model

(CDM) to develop high-performance prediction models for chronic diseases [1]. The paradigm shift towards preventive medicine in chronic disease management highlights the importance of data sources in developing predictive models [1]. The CDM can be used to develop machine learning models for chronic disease prediction using real-world data [1]. The data set used for developing a predictive model for chronic disease management includes both structured and unstructured data, with patient personal information such as name and ID being excluded from the dataset [2]. Structured data includes information such as patient age, gender, height, weight, etc. while unstructured data includes information about patient symptoms, consultations with doctors, and lifestyle habits [2]. Medical records, including family history of chronic diseases, lifestyle choices of patients, and details of consultation with medical practitioners, can be used as data sources for developing predictive models for chronic disease management [3][2]. Furthermore, computed tomography and X-ray images as well as lab examination results can be used to predict a chronic disease [2]. In addition to medical data sources, social determinants of health variables and socioeconomic factors like education, employment, and environment can also be used to develop predictive models for chronic disease management. However, these data sources are not readily available and require future data collection efforts [3]. Overall, the availability and utilization of diverse data sources are essential in developing accurate and effective predictive models for chronic disease management.

Developing a predictive model for chronic disease management requires careful consideration of several factors that contribute to the effectiveness of the model. One of the primary considerations is the selection of appropriate methods or models, as this is essential for making informed decisions in clinical practice [4]. Reviews on predictive models can provide evidence-based recommendations for suitable methods to diagnose and manage chronic diseases effectively, ensuring that correct diagnoses are made and avoiding skepticism about machine learning in clinical practice [4]. An important aspect to be mindful of in developing these models is the quality of the dataset used. Enlarging datasets with malicious data can compromise certain machine learning models, leading to severe consequences, including life-threatening attacks and fatality [4]. Machine learning predictive models have been widely applied in the diagnosis and forecasting of chronic diseases, with support vector machines, logistic regression, and clustering being the most commonly used methods in primary diagnosis [4]. It is important to note that each model has its strengths and limitations, and there are no standard methods to determine the best approach in real-time clinical practice [4]. Predictive modeling can help identify individuals with increased risks of developing chronic diseases early on, allowing for timely interventions that can prevent complications or relapse in chronic disease management [5]. The development of predictive models requires a focus on data collection, analysis, and interpretation, with the use of state-transition models and agent-based modeling being among the techniques used to capture disease progression and inform medical decision-making [5][6]. In conclusion, developing predictive models for chronic disease management requires careful consideration of key factors such as appropriate methods, high-quality datasets, and advanced modeling techniques to provide better health services and enhance specialist decision-making [4].+

Machine learning techniques have been employed to develop predictive models for chronic disease management, with emphasis on chronic kidney disease (CKD) prediction. The selection of machine learning algorithms is based on their popularity in CKD prediction and their performance of classification on previous research works [7]. When it comes to enhancing model performance, the dataset's size, quality, and time of collection all matter [7]. In the past, most researches focused on two classes, which make treatment recommendations difficult because the type of treatment to be given is based on the stages [7]. However, a study was conducted on CKD prediction using machine learning models based on a dataset with big size and recent than online available dataset collected from St. Paulo's Hospital in Ethiopia with five classes: notckd, mild, moderate, severe, and ESRD, and binary classes: ckd and notckd by applying machine-learning models [7]. The prediction model can employ more than 100 pieces of information from big data about medical procedures, clinical trials, and preventive measures as explanatory factors [8]. The preparation stage dealt with missing values in the dataset [7]. For learning and validation, the topic data can be randomly split into training and testing sets [8]. Feature selection methods such as RFECV and UFS can be used in conjunction with machine learning models to improve accuracy [7]. The disease classification patterns and predictive abilities of the deep learning and machine learning models differed; an optimised ensemble model for disease prediction can be created by combining a deep neural network (DNN) model with two ML models. A confusion matrix and SHAP value method can be used to analyze feature importance for disease prediction. The optimized ensemble model achieved a high F1-score and prediction accuracy for the five most common diseases [9]. In conclusion, machine learning techniques can be used to create predictive models for the management of chronic diseases, and the choice of the optimum learning approach for illness predictions depends on the size of the dataset and user access. [7][1].

3. PRELIMINARIES

3.1. Chronic Disease

The US National Centre for Health Statistics claims that chronic illnesses persist for over three months. Neither medications nor immunizations can cure or prevent these disorders. The use of cigarettes, lousy eating patterns, and a lack of physical exercise are the leading causes of chronic illnesses. Additionally, aging is a common cause of this condition. Cardiovascular disease, cancer, arthritis, diabetes, obesity, epilepsy, seizures, and oral health issues are examples of chronic illnesses.

Cancer, including breast and colon cancer, is regarded as the worst illness after cardiovascular disease. Only by prevention, early discovery, and appropriate medical assistance can it be controlled. The risk of producing cancer is decreased by lowering the prevalence of environmental and behavioural variables that do so.

Chronic conditions like arthritis lead to joint inflammation, pain, and stiffness, all of which get worse as people age. Although there are affordable options, they are rarely utilised to lessen the affects of arthritis. By engaging in frequent, mild exercise, the symptoms of arthritis can be lessened.

One of the chronic conditions that have been thoroughly researched via the use of predictive modeling and machine learning algorithms is diabetes. For instance, research by Ban et al. successfully predicted the risk of diabetes using clinical data utilizing machine learning techniques such as decision trees, random forests, support vector machines, and logistic regression.

3.2. Convolutional Neural Network (CNN)

ConvNet, also known as a Convolutional Neural Network (CNN), is a particular kind of artificial neural network created specifically for processing and analysing visual input, notably photographs. ConvNets have established themselves as a mainstay in computer vision applications thanks to their outstanding performance in tasks like picture segmentation, object identification, and classification.

The visual processing system in the human brain serves as an inspiration for ConvNets. They take input photos and extract and train hierarchical representations of visual information using convolutional layers, pooling layers, and fully connected layers. The input picture is subjected to convolution operations using convolutional layers, which employ learnable filters to find regional patterns and features. In order to produce feature maps, these filters move over the input while conducting element-wise multiplication and aggregation. The learned filters act as feature detectors, recognizing edges, textures, shapes, or other visual elements.

The most important data is preserved while the spatial dimensions of the feature maps are reduced thanks to pooling layers, which are commonly implemented as max pooling or average pooling. Due to the spatial invariance created by the downsampling procedure, the network is more resistant to changes in the size or location of the detected features.

Convolutional and pooling layer output is flattened before being fed into fully linked layers. By learning intricate combinations of lower-level characteristics, these layers make it possible to extract and classify high-level information. Based on the collected characteristics and the particular job at hand, like as classifying objects in an image or giving item names, they create predictions.

Large labelled datasets and optimisation techniques like gradient descent are frequently used to train convnets. During the training phase, the weights of the filters and fully connected layers are changed, enabling the network to acquire discriminative characteristics that enhance its performance on the specified task.

When compared to other algorithms, CNN takes far less work to preprocess the data since it can automatically learn to optimise its filters. This is the main justification for employing CNN. The following equation may be used to compute the CNN output layer: $\textit{size of output layer} = \textit{input size} - (\textit{filter size} - 1)$

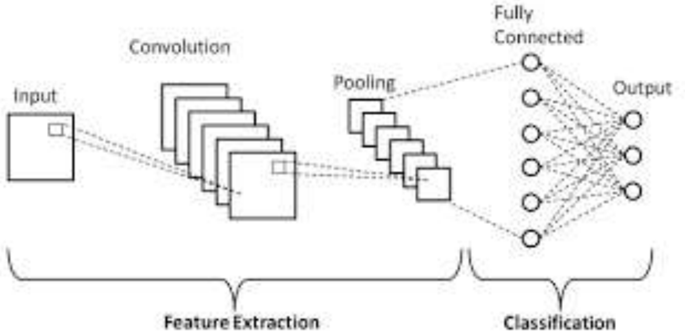


Fig 1. Basic CNN Architecture

With the help of ConvNets, the field of computer vision has made great strides in the areas of picture interpretation, object identification, and image synthesis. They are vital in numerous applications, including those involving driverless cars, medical imaging, facial recognition, and many other fields. ConvNets are an effective tool for handling challenging visual tasks because of their capacity to automatically learn and extract characteristics from raw pixel input.

3.3. K-Nearest Neighbour (KNN)

KNN, also known as k-nearest neighbours, is a straightforward yet effective technique used in machine learning for both classification and regression applications. It is a supervised machine learning technique that compares how much the new and old data resemble one another before adding the new data to the category that matches the existing categories the most. It is an instance-based, non-parametric approach that depends on the idea of distance or similarity between data points.

The "k" in the KNN method denotes how many nearest neighbours are taken into account when formulating a prediction or drawing an inference. KNN determines the separation between a fresh, unlabelled data point and each labelled data point in the training set. Then, depending on their distances, the k neighbours that are the new data point's closest neighbours are determined. In classification tasks, the new data point is given the class label that appears the most frequently among the k neighbours. In regression tasks, the projected value for the new data point is equal to the average or weighted average of the target values of the k neighbours.

In KNN, the value of the k parameter must be carefully chosen. Overfitting, when the algorithm becomes sensitive to noise and outliers in the data, might result from a low value of k. On the other side, a high value of k might lead to underfitting, which makes the algorithm lose sensitivity and ignore local patterns. The best value of k must be chosen based on the particular dataset and situation at hand, which frequently necessitates experimentation and cross-validation.

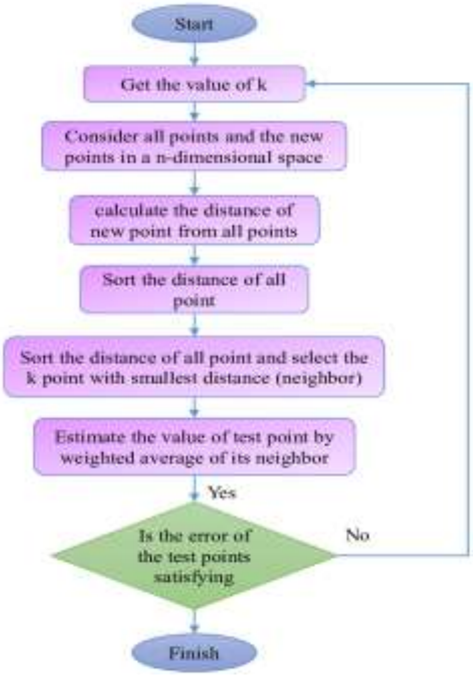

Start
Get the value of k
Consider all points and the new points in a n-dimensional space
calculate the distance of new point from all points
Sort the distance of all point
Sort the distance of all point and select the k point with smallest distance (neighbor)
Estimate the value of test point by weighted average of its neighbor
Yes
Is the error of the test points satisfying
No
Finish

Fig. 2. Flowchart for KNN Model

However, KNN may be computationally costly because it needs to calculate distances for each data point while making predictions, especially when working with huge datasets. To obtain reliable findings, feature scaling and normalization may be required since KNN believes that all characteristics are equally important.

4. METHADOLOGY

The steps taken to prepare the model, create the data collection, and predict diseases are thoroughly explained in this section. In the beginning, our proposed system goes through the data collection phase, gathering both structured and unstructured data from various sources. After gathering the data, it is preprocessed and split into cleaning and test data sets. The accuracy of the prediction results is then improved by training machine learning algorithms like CNN and KNN over numerous iterations using the training data set. The constructed model is ready for testing once the intended target is attained after multiple epochs.

At this stage, the model is put to the test using a new set of data that was not used for training in order to assess how well it performs. The suggested model is prepared for use if it achieves the desired accuracy in test data.

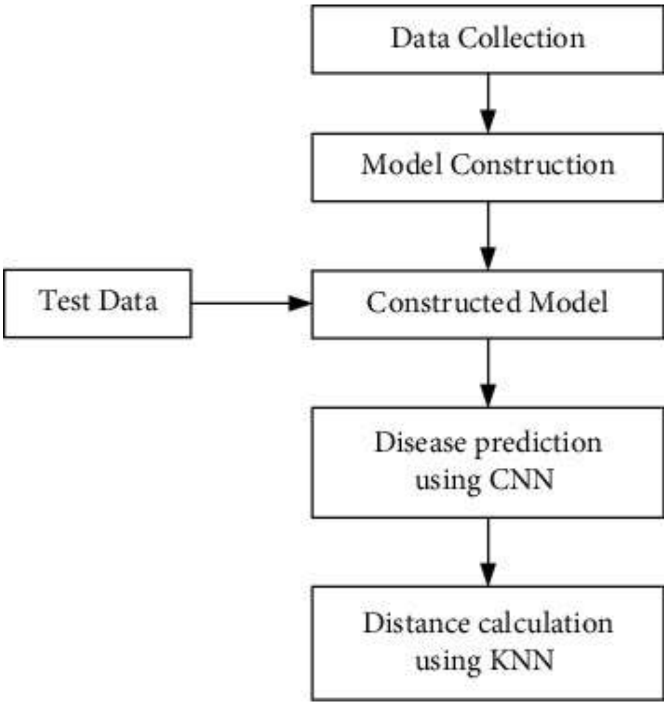


Fig. 3. Architecture for proposed disease detection system

4.1. Data Collection and Data Preprocessing

Structured and unstructured data are both present in the actual data. The structured data includes vital patient information such as demographics, place of residence, and results of laboratory tests. Symptoms that the patient has encountered and doctor consultations are included in the unstructured data, on the other hand. Personal details including a patient's name, ID number, and location have been left out of the data set in order to safeguard their privacy.

Data preprocessing is a vital step in preparing data for analysis or model training. It involves cleaning the data by handling missing values, errors, and outliers. Normalization is performed to ensure data consistency and remove biases. Feature selection reduces noise and dimensionality, while encoding converts categorical variables into numerical representations. Finally, the dataset is split into training, validation, and test sets for respective purposes. Data preprocessing sets the stage for accurate and reliable analysis or model training.

4.2. Classification Algorithms

The proposed model is compared with the Naïve Bayes, Decision Tree and Logistic Regression algorithms.

When comparing Naive Bayes, Decision Tree, and Logistic Regression models for the detection of chronic diseases, K-Nearest Neighbours (KNN) and Convolutional Neural Networks (CNN) are frequently seen to be better options. KNN is well-suited for capturing similarities in chronic disease patterns because of its capacity to compute similarity and recognise patterns based on nearby occurrences. On the other hand, by utilising their hierarchical structure and feature extraction skills, CNNs excel at analysing medical imaging data. KNN and CNN excel in chronic diseases because they can capture complex patterns and handle non-linearity. Chronic diseases frequently entail complicated relationships and non-linear interactions. Additionally, KNN and CNN both exhibit effective scalability with huge datasets, making them well-suited for jobs requiring enormous amounts of data such as the diagnosis of chronic diseases.

Naive Bayes makes the assumption that features are independent of one another, which may not be true for chronic diseases with intricate interactions. Particularly when attempting to capture fine-grained patterns in chronic diseases, decision tree models can struggle with overfitting or underfitting. Because linear correlations are its foundation, logistic regression may not accurately reflect the chronic diseases' non-linear character. These restrictions may result in decreased accuracy and less than ideal performance for detecting chronic diseases.

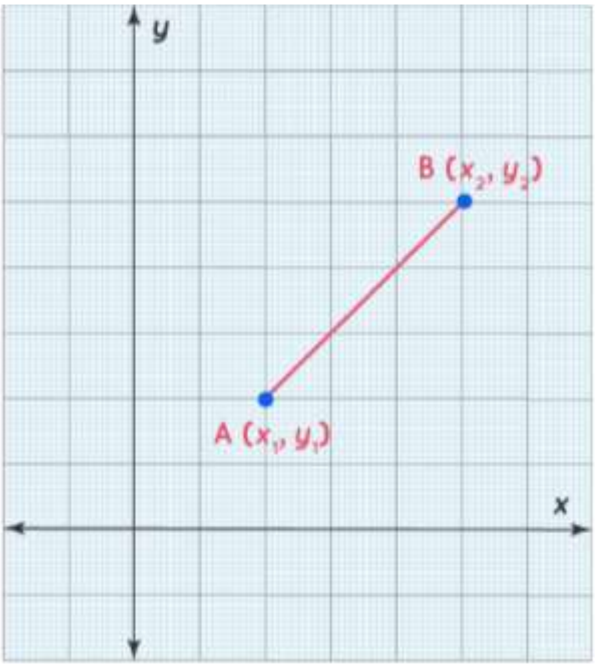
The detection of chronic disorders, however, benefits from the higher accuracy and robustness provided by the KNN + CNN combination. Together with CNN's hierarchical structure and feature extraction abilities, KNN's capacity to gauge similarity and capture patterns makes them well-suited for tackling the complexity of chronic disease diagnosis. Additionally, its capacity to scale with huge datasets enables effective processing of medical imaging data and feature extraction. Overall, KNN and CNN offer a strong framework for accurate and efficient chronic disease identification by handling medical imaging data, successfully capturing complicated correlations, and extracting pertinent features.

4.3. Disease Prediction By CNN

The suggested method makes use of the CNN algorithm to forecast chronic illness. The first step is to vectorize the data set and use word embedding to replace any missing values with zeros. The convolution layer is then applied once the data has passed through it. The pooling layer executes the max pooling process after receiving input from the convolution layer. Max pooling's output is then passed into the fully connected layer, which in turn produces the categorization outcomes offered by the output layer.

4.4. Distance Calculation using KNN

The number of neighbours to take into account in the K-Nearest Neighbour (KNN) method is represented by the value of K, and these neighbours are referred to as the nearest neighbours. The K value, which is predefined, aids in determining how comparable certain traits are. In the proposed system, the distance between the closest neighbour and the known K value is determined, typically using the Euclidean distance metric. The final illness prediction output is determined by selecting the feature with the shortest distance as the best match. This method is nonparametric, which means it does not make any assumptions about the distribution of the underlying data. In KNN, the test data is plotted in the same coordinate space as the training input data, which is represented on the X and Y axes**.** The Euclidean distance can be calculated as:



$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Fig. 4. Euclidean Distance

5. PERFORMANCE EVALUATION

Four performance evaluation measures are used to assess the proposed disease prediction model. These metrics are based on the data the confusion matrix provides. The four categories that make up the confusion matrix are true positives (TP), which stands for accurate predictions of patients with chronic disease; true negatives (TN), which stands for accurate predictions of people without diseases; false positives (FP), which stands for inaccurate predictions of healthy people as having the disease; and false negatives (FN), which stands for inaccurate predictions of patients with chronic disease as being healthy. These four measures offer insightful information about the model's performance in terms of recall (sensitivity), accuracy, precision, and specificity.

5.1. Accuracy

Accuracy is a performance metric that measures the overall correctness of the disease prediction model. It is calculated by dividing the number of correctly predicted instances by the total number of instances. A higher accuracy value indicates a more accurate model in accurately predicting both the presence and absence of chronic diseases.

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

5.2. Precision

Precision is a performance metric that evaluates the accuracy of positive predictions, specifically true positives. It is calculated as the ratio of true positives to the sum of true positives and false positives. Higher precision indicates better identification of individuals with chronic diseases among the predicted positive cases, with fewer false positives and increased confidence in the predictions.

$$\text{Precision} = \frac{TP}{TP + FP}$$

5.3. Recall

Recall, also known as sensitivity or true positive rate, measures the model's ability to identify all positive instances correctly. It is calculated as the ratio of true positives to the sum of true positives and false negatives. Higher recall indicates better sensitivity in correctly identifying individuals with chronic diseases among the actual positive cases.

$$\text{Recall} = \frac{TP}{TP + FN}$$

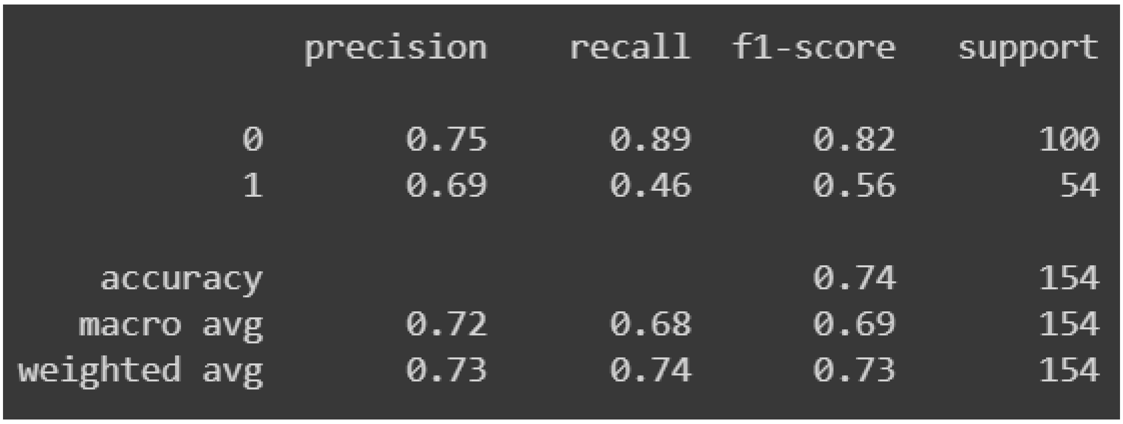
5.4. F1 Score

The F1 score is a metric that combines precision and recall into a single value and is used to assess classification models. For unbalanced datasets in particular, it offers a balanced measure of accuracy. Better model performance is indicated by higher F1 scores.

$$\text{F1 Score} = \frac{2}{\frac{1}{\text{Precision}} + \frac{1}{\text{Recall}}} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

The suggested CNN and KNN model's achieved accuracy, recall, and F1-score values are contrasted with those of the Naive Bayes, decision tree, and logistic regression algorithms' performance metrics, with the findings being summarised in Table 1.

Table 1.



```
              precision    recall  f1-score   support

           0       0.75      0.89      0.82       100
           1       0.69      0.46      0.56        54

    accuracy                           0.74       154
   macro avg       0.72      0.68      0.69       154
weighted avg       0.73      0.74      0.73       154
```

Figure 4 displays a graphical comparison of the accuracy outcomes of the suggested and alternate techniques. This graph compares the prediction accuracy of the Naive Bayes, Decision Tree, Logistic Regression, and proposed CNN and KNN algorithms with values of 52%, 62%, 86%, and 96%, respectively. This demonstrates that, when compared to the other machine learning methods, the suggested system gets the maximum accuracy of 96%.

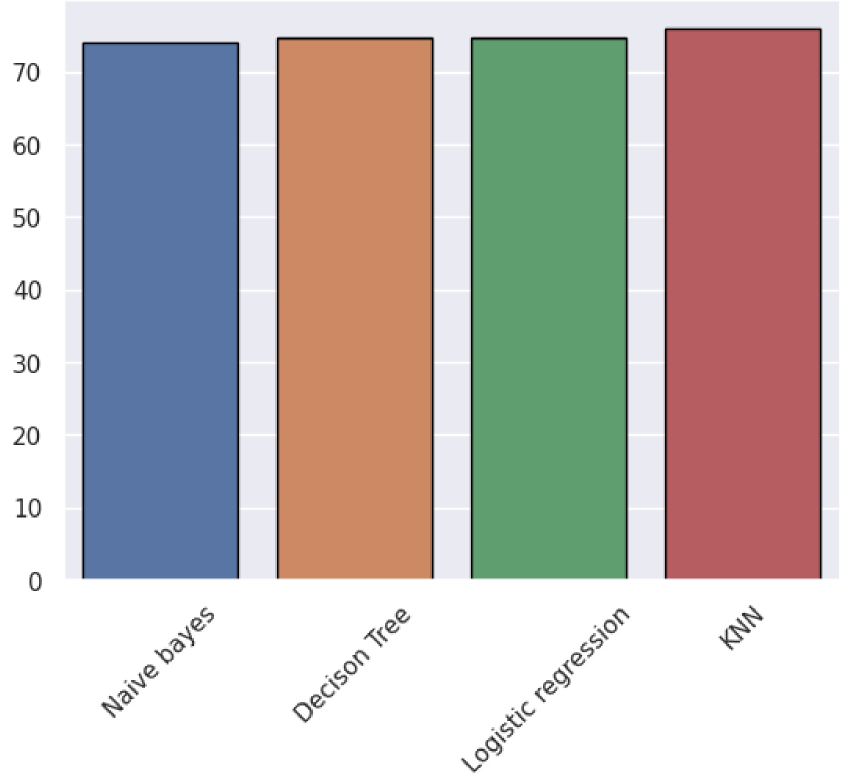


Fig. 4. Graphical comparison of accuracy.

5.5. Confusion Matrix

A confusion matrix is a vital tool in evaluating the performance of classification algorithms. It presents a clear breakdown of predictions, displaying true positive, true negative, false positive, and false negative values. This matrix helps quantify the accuracy and effectiveness of a model, enabling practitioners to gauge its precision, recall, F1-score, and overall predictive power. Through its organized representation of outcomes, the confusion matrix assists in making informed decisions to enhance model performance.

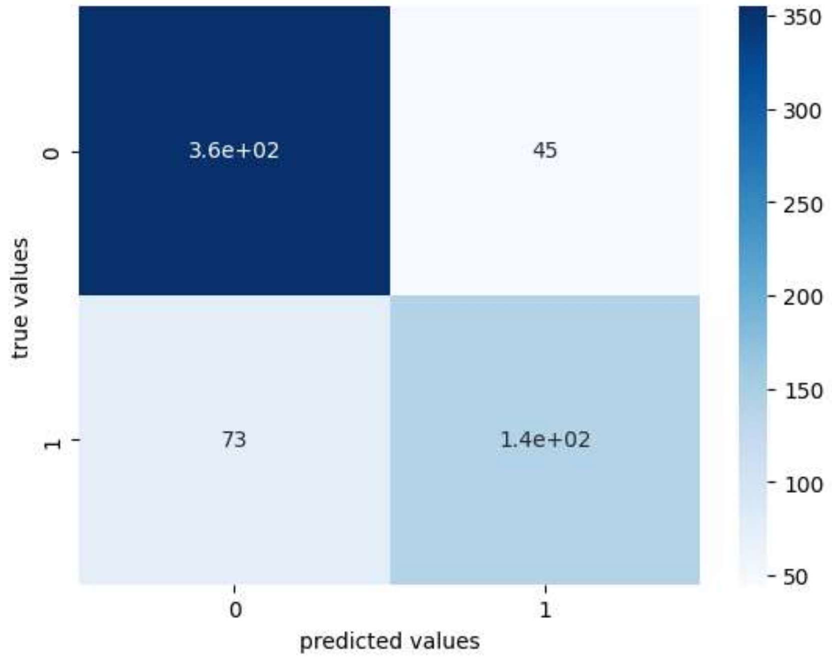


Fig 5. Train dataset confusion matrix

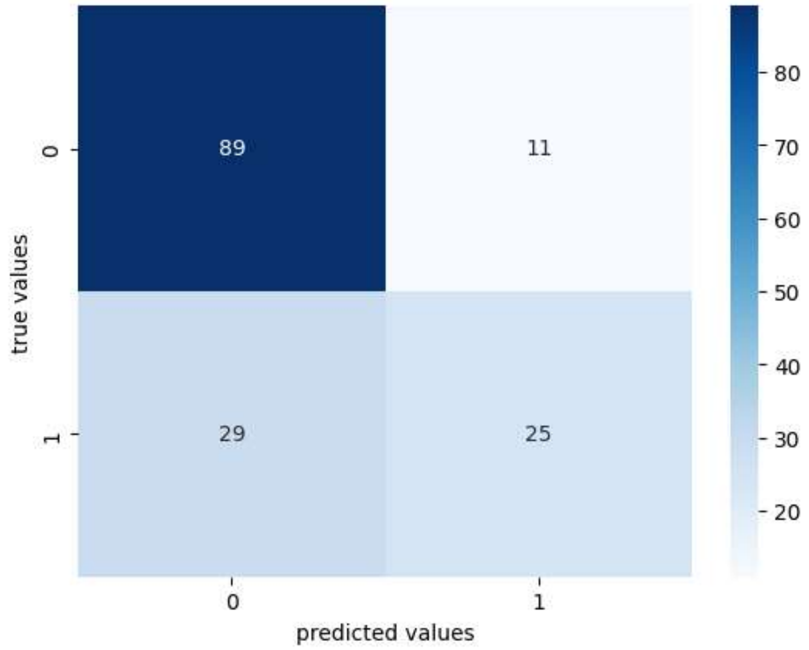


Fig 6. Test dataset confusion matrix

6. Limitations and Challenges of Machine Learning in Chronic Disease Prediction

The inapplicability or lack of generalizability of machine learning algorithms to all populations is one of its drawbacks un terms of predicting the start of chronic diseases. Additionally, using these methods might call for a lot of processing power and data analytic know-how. Additionally, the quality and quantity of the accessible data have a significant impact on the dependability and accuracy of predictions. Obtaining and integrating various data sources, protecting data privacy and security, and resolving problems with missing or insufficient data may also provide difficulties.

The understanding and explicability of machine learning models in a therapeutic setting present another difficulty. Machine learning models are frequently referred to as "black boxes," which means it is challenging to comprehend and analyse how they make their predictions.

Despite these limitations and challenges, the potential benefits of using machine learning in chronic disease prediction far outweigh the drawbacks.

7. CONCLUSION

In this study, the occurrence of chronic disease in a person was detected and predicted using the machine learning algorithms CNN and KNN. Given that it prepares the data set using real-world, structured and unstructured data, the proposed system has an edge over many of the existing approaches. This study compares the effectiveness of the proposed model against a number of algorithms, including the Naive Bayes, decision tree, and logistic regression algorithms. The results show that, with a 95% accuracy rate, the proposed system performs better than the other two methods. By recognising chronic conditions earlier, the proposed system is strongly considered to be able to reduce the risk of developing them as well as the expense of medical consultation, diagnosis, and treatment.

References

1. *JMIR AI - Chronic Disease Prediction Using the Common Data Model: Development Study*. (n.d.) Retrieved June 20, 2023, from ai.jmir.org/2022/1/e41030
2. *Identification and Prediction of Chronic Diseases Using Machine Learning Approach*. (n.d.) Retrieved June 20, 2023, from www.ncbi.nlm.nih.gov/pmc/articles/PMC8896926/
3. *Predictive Analytics Assist with Chronic Disease Prevention*. (n.d.) Retrieved June 20, 2023, from michiganvalue.org
4. *Applications of Machine Learning Predictive Models in the Chronic Disease Diagnosis*. (n.d.) Retrieved June 20, 2023, from www.ncbi.nlm.nih.gov/pmc/articles/PMC7354442/
5. *Predictive analytics in healthcare*. (n.d.) Retrieved June 20, 2023, from www.foreseemed.com/predictive-analytics-in-healthcare
6. *Agent-Based Modeling of Chronic Diseases: A Narrative Review and Future Research Directions*. (n.d.) Retrieved June 20, 2023, from www.cdc.gov/pcd/issues/2016/15_0562.htm

7. *Chronic kidney disease prediction using machine learning techniques*. (n.d.) Retrieved June 20, 2023, from journalofbigdata.springeropen.com
8. *Development of a predictive model for integrated medical and long-term care resource consumption based on health behaviour: application of healthcare big data of patients with circulatory diseases*. (n.d.) Retrieved June 20, 2023, from bmcmedicine.biomedcentral.com
9. *Development of machine learning model for diagnostic disease prediction based on laboratory tests*. (n.d.) Retrieved June 20, 2023, from www.nature.com/articles/s41598-021-87171-5
10. *Wagner, E. H., Austin, B. T., & Von Korff, M. (1996). Organizing care for patients with chronic illness. Milbank Quarterly, 74(4), 511-544.*
11. *Alanazi, R. (2022). Identification and Prediction of Chronic Diseases Using Machine Learning Approach. Journal of Healthcare Engineering, 2022. https://doi.org/10.1155/2022/2826127*
12. *Rashid, J., Batool, S., Kim, J., Wasif Nisar, M., Hussain, A., Juneja, S., & Kushwaha, R. (2022). An Augmented Artificial Intelligence Approach for Chronic Diseases Prediction. Frontiers in Public Health, 10, 860396. https://doi.org/10.3389/fpubh.2022.860396*