iJRASET
International Journal For Research in Applied Science and Engineering Technology

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 11 **Issue:** VI **Month of publication:** June 2023

**DOI:** https://doi.org/10.22214/ijraset.2023.54499

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 11 Issue VI Jun 2023- Available at www.ijraset.com*

# Multi Disease Detection Using Deep Learning

Prof. R.Y. Totare[1], Ishwari Ambre[2], Samruddhi Deore[3], Aryan Ingle[4], Akanksha Vetal[5]
*[1]Professor, [2, 3, 4,5]Student, , Department of Information Technology, AISSMS IOIT, SPPU, Pune*

***Abstract:*** ***Chronic diseases are the main reason for increase in mortality rate. Heart disease, cancer, diabetes, stroke, and arthritis all are chronic diseases. NCDs have a more mortality rate than other diseases. Each year, approximately 5.8 million Indians die from NCDs; globally, 41 million die from chronic diseases. Investment and prevention are two important needs for chronic diseases. It is important to discover the solution for chronic diseases. The management of these types of illnesses includes early detection, patient care, and related services. Because of the explosion of medical data, data management will become a difficult challenge. Deep learning [9] is important in big data fields such as the medical industry. Deep learning is a advancement of machine learning that is capable of performing different tasks. Chronic diseases are diagnosed using different technologies, and that is why deep learning is used to provide the best treatment to patients, resulting in a low mortality rate.***
***Keywords: Chronic, Diagnosis, Detection, Healthcare, Prediction, Deep Learning, Machine Learning.***

## I. INTRODUCTION

As time passes, environmental conditions, societal progress, and people's lifestyles change, all of which have a significant impact on health. Chronic diseases have a more effect all over the world. Looking only at cardiovascular diseases, the annual death rate in India is 27%. The main purpose of disease detection and stage prediction are basically to reduce mortality rate by detecting diseases at early stages using Deep Learning [5]. A large number of data are generated in the medical field [6]. But in the medical industry, the complexity level of data is different compared to other fields. They are very vulnerable to impact.Existing deep learning systems only focuses on one disease detection per testing. There is a need for one common system that will work simultaneously on many diseases. We are giving a system that will solve this problem.

In this proposed model, we examine chronic diseases such as heart disease, cancer, diabetes, etc. To perform multiple disease prediction,we using machine learning algorithms [1]. In this system, parameters are added while analyzing the diseases. Various disease detection is done by machine learning algorithms, for example, Lung Cancer [2]. As we said parameters to the system, we could detect the disease efficiently, at early stages [3] and the accuracy of the diagnosis also increased. This system is an efficient alternative to manual diagnosis techniques [8]. Doctors can cross-verify the test results .The cost of testing for NCDs can be reduced by using this system. The experience of doctors will also increase because of this model.

## II. SYSTEM ARCHITECTURE

System uses machine learning and deep learning algorithms for disease prediction[7]. In Multi Disease Prediction System using Deep learning, several phases are shown in Fig. 1. The initial step is to collect the patient's data. After we import the dataset, on each input pre-processing will occur. At the end of visualization, data pre-processing starts. In this step, the system checks for outliers and missing values and scales the dataset.

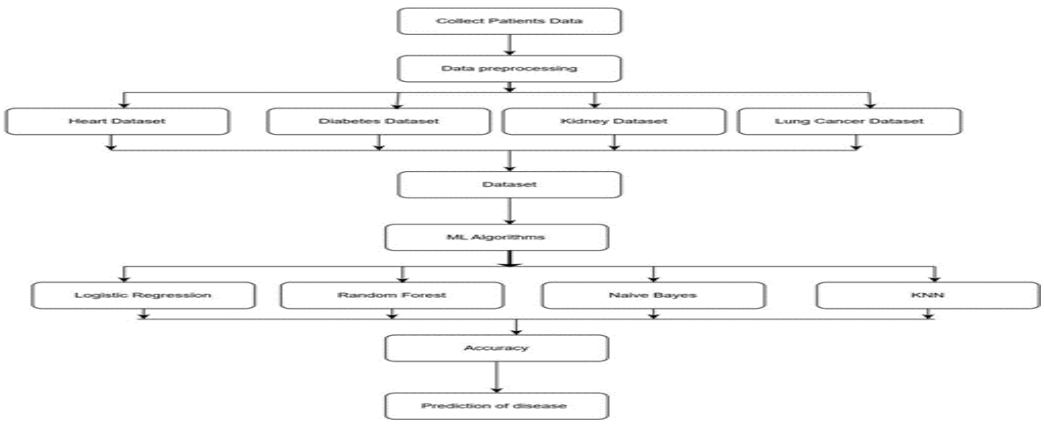



Fig 1. System Architecture

Following the completion of data pre-processing, the system used KNN, random forest, naive Bayes, and logistic regression algorithms. Using a test dataset, this system selects the best algorithm for correct accuracy for each NCD. After that, the system integrated the file with the Django framework. This helps export the model to the web.

## III. METHODOLOGY/ALGORITHM DETAILS

For the detection and prediction of multi-disease using deep learning our system uses various methodologies [4]. Data collection and data analysis, Algorithm techniques, evaluation of accuracies, and model comparison are the steps of this system.

### *A. Data Collection And Analysis*

Data collection and analysis include datasets, training data, testing data, and balanced data. For the dataset, we collect data from Kaggle. Training data is an important aspect of the deep learning model. Training data is used for the accurate prediction of diseases. Testing data is used for performance evaluation collection.

### *B. Algorithm Techniques*

1) *CNN:* CNN is mostly used for image recognition and also for the classification of images.CNN is a Convolutional Neural Network.This model uses deep learning to find objects in an image with the help of convolutional neural networks. CNN [10] extracts different features of the image. CNN takes in input images. CNN have different layers. They are used to extract features of image input.CNN is made up of three layers: CONV, POOL, and fully connected (FC).The convolutional layer extracts different features from the input image. Next, the pooling layer is used to reduce the size of pixel block representation and the amount of computation in the network. The fully connected layer is the last layer of CNN. Fig.1.2 Shows the CNN architecture.

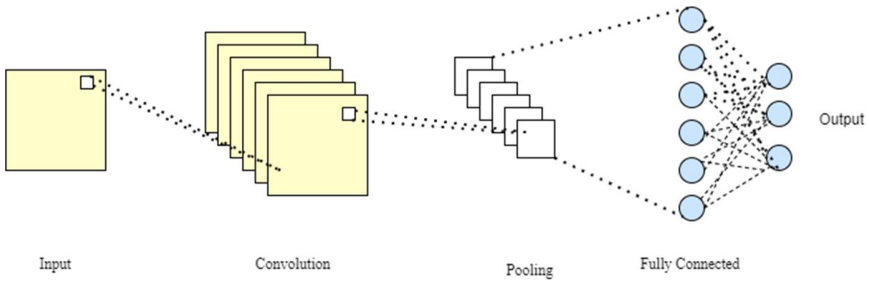



Fig 2. CNN Architecture

2) *Logistic regression:* Logistic regression is a machine-learning technique.It is a binary classification technique. Logistic regression is used for finding objects in input image.It predicts a dependent variable based on more than one set of independent variables to get the output. It helps find the categorical dependent variables. The outcome value is between 0 and 1. This regression is used to solve regression problems. Here, logistic regression is used for solving classification problems. Fig.3. Shows, Logistic regression models the data using the sigmoid function

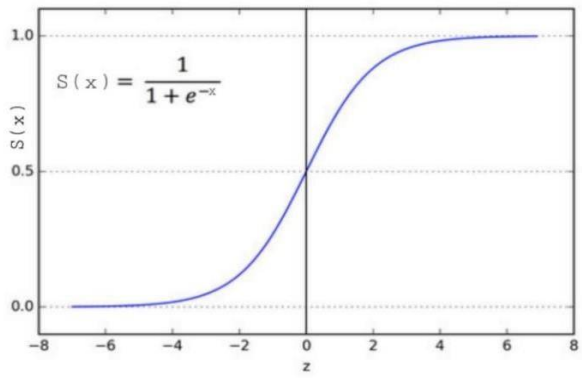



Fig.3 Logistic Regression

3) *Random Forest:* Random Forest is the most common machine learning algorithm which is widely used. This algorithm belongs to the supervised learning algorithm. Random Forest constructs different decision trees. Decision trees suffer from low bias and high variance. This algorithm is used to convert high variance to low variance. Below is Fig .4 explains the working of the Random Forest algorithm.

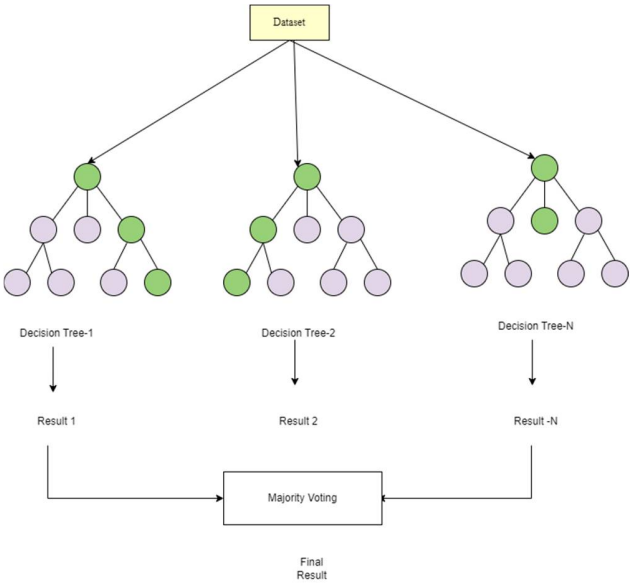



Fig. 4 Random Forest

*C. Calculate the Accuracies of the Deep learning models*

The confusion matrix is used to find the performance of the model. This confusion matrix compares the actual value with the value which is get by deep learning models. To find the performance evaluation in this system, we denote TP, TN, Fp, and FNias. We can get these measurements: Precision, recall, and f1 measures as follows:

Accuracy= (True Positive+ True Negative) / (Total Sample Size)

Precision = True Positive/ (True Positive+False Positive)

Recall = True Positive/ (True Positive + False Positive)

F1 Score = 2 * (Precision * Recall) / (Precision + Recall)

*D. Model Comparison*

For diabetes disease detection and prediction this system uses models and algorithms such as KNN, Logistic Regression, SVM, RF, XgBoost, Decision Tree Classifier, and Gradient Boosting Classifier.

Performance evaluation for diabetes disease detection and prediction is different for different models. Fig. 5 shows the Performance evaluation for diabetes disease prediction.

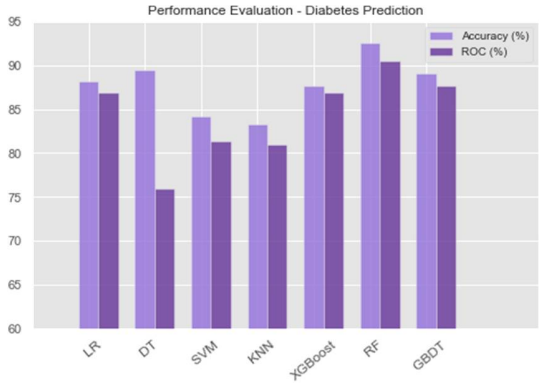



Fig.5 Performance evaluation for diabetes

## IV. RESULT AND EXPERIMENTS

This model proposes a deep learning-based multiple- disease prediction model for chronic diseases with higher accuracy. To solve accuracy issues in the diagnosis of diabetes with accurate stage predictions.For experimental results, the "diabetes" data set from www.kaggle.com is obtained. Various evaluation measures like accuracy, precision, recall, and f1-score are used to check the effectiveness of the system. Fig.6 shows the confusion matrix of diabetes obtained.

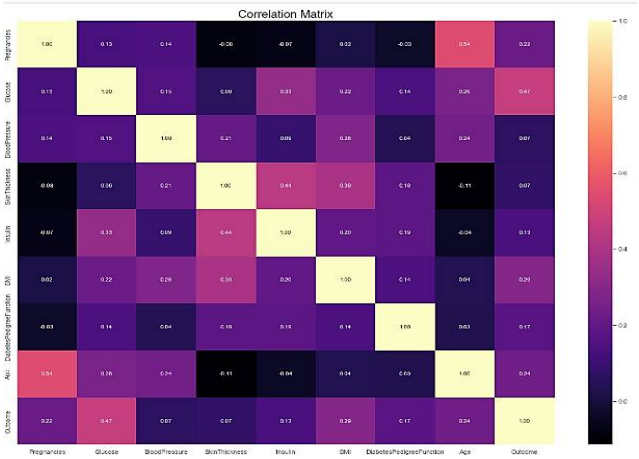



Fig.6 Confusion Matrix

Table 1 shows the precision, recall, and f-measure during the validation of the system.

| | Precision | Recall | f1-score | Support |
|---|---|---|---|---|
| 0 | 0.95 | 0.94 | 0.94 | 147 |
| 1 | 0.89 | 0.90 | 0.90 | 81 |
| Accuracy | 0.93 | 0.93 | 0.93 | 228 |

## V. CONCLUSIONS

The primary goal of this system is to detect and predict chronic diseases at an early stage to reduce mortality rates. Given the explosion of medical data and its complexity, deep learning plays a major role in disease detection. This is used to create a more precise prediction of diseases. This model will help doctors cross-verify the test results provided by the labs. This system can be used by doctors to enhance their experience with diseases. By using this model, we will be able to reduce the cost of the tests that need to be carried out for the detection of chronic disease

## REFERENCES

[1] D. Dahiwade, G.Patle, and E. Meshram, "Designing disease prediction model using machine learning approach," 2019 3rd International Conference on Computing Methodologies and Communication 2019, pp 1211-1215, doi:10.1109/ICCMC.2019.8819782.
[2] K. Karthick, S. Rajakumar, N. Selvanathan, U.K, B. Saravanan, M. Murali, and B. Dhiyanesh, "Analysis of Lung Cancer Detection Based on the Machine Learning Algorithm and IOT," 2021 6th International Conference on Communication and Electronic System (ICCES), 2021.9489084.
[3] R. Katarya and P. Srinivas, "Predicting Heart Disease at Early Stages using Machine Learning: ASurvey,'' 2020 International Conference on Electronics and Communication System (ICESC), 2020, pp.302-305, doi: 10.1109/ICESC48915.2020.9155586.
[4] G. Sruthi, C. L. Ram, M. K. Singh, N. Malhotra, and N. Sharma, " Cancer Prediction Using Machine Learning," 2022 2nd International Conference on Innovative Practices in Technology And Management (ICIPTM),2022, pp.217-221,doi:10.1109/ICIPTM54933.2022.9754059.
[5] B. Lalithadevi and S. Krishnaveni, "Efficient Disease Risk Prediction based on Deep Learning Approach," 2022 6th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2022,pp. 1197-1204, doi:10.1109/ICCMC53470.2022.9753851.
[6] Prof. Reshma Totare, Anjali Singh, Vishakha Swamy, Tejaswini Yadav "Advanced Database Update and Security Using Markov Decision Policy," 2021.

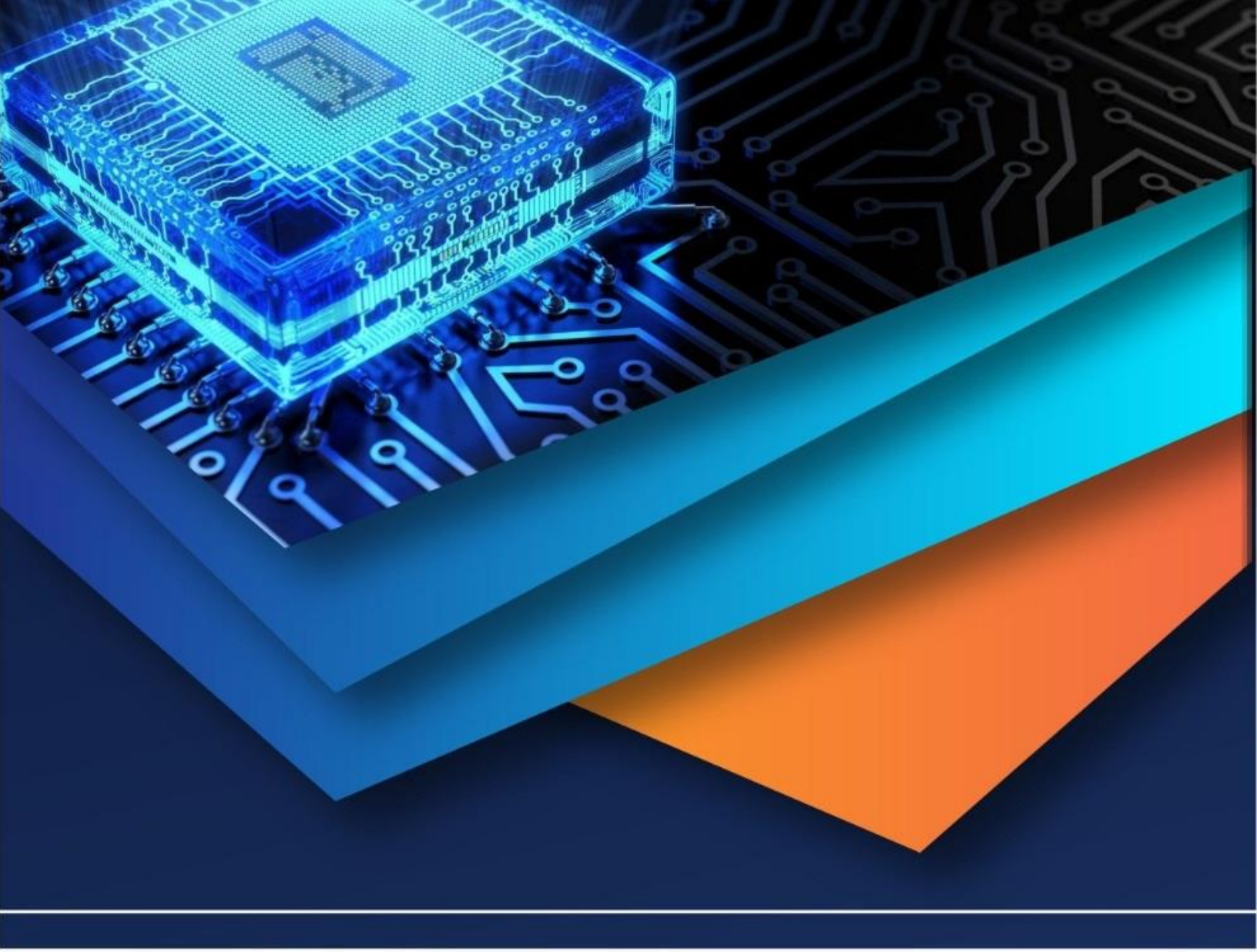
[7] K. Vayadande et al., "Heart disease prediction using machine learning and deep learning algorithms," 2022 International Conference on Computational Intelligence and Sustainable Engineering Solutions (CISES), Greater Noida, India, 2022, pp. 393-401, doi: 10.1109/CISES54857.2022.9844406.

[8] P. M, Y. B, S. S, and S. M, "Cardiovascular disease prediction using deep learning," 2022 6th International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 2022, pp. 1399-1404, doi: 10.1109/ICOEI53556.2022.9777135.

[9] M. Chen, Y. Hao, K. Hwang, L. Wang, and L. Wang, "Disease prediction by machine learning Over Big Data From Healthcare Communities," in IEEE Access, vol. 5, pp. 8869-8879, 2017, doi: 10.1109/ACCESS.2017.2694446.

[10] A. S. N. Zainab, I. Soesanti and D. R. Utomo, "Detection of COVID-19 using CNN's Deep Learning Method: Review," 2022 4th International Conference on Biomedical Engineering (IBIOMED), Yogyakarta, Indonesia, 2022, pp. 59-64, doi: 10.1109/IBIOMED56408.2022.9988533.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY