JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Implementation Of Healthcure Using Machine Learning

Poonam V. Jagtap, Pratidnya B. Shikhare, Pratiksha D. Shelar and Anubhav D. Bhat

Project Guide – Mr. Ashish K. Bhise

Department Of Computer Science Engineering

Marathwada Mitra Mandal Institute of Technology(MMIT), Lohegaon-411047, India

Abstract **: Nowadays people face different diseases because of the state of the climate and their living habits. While earlier prediction of disease becomes an essential task. But it becomes too difficult for the doctor to determine precisely based on symptoms. Data mining plays an important part in predicting the disease in overcoming this. Medical science has a substantial increase in data per annum. We proposed general prediction of the disease based on the patient 's symptoms. For the prediction of disease, we use machine learning algorithm CNN for accurate prediction of disease. Dataset of disease symptoms needed for prediction of disease. It is difficult to handle the huge amount of data of the patients. It is easier to handle this data through Big Data Analytics. Peoples are being affected by pollution and irregular food habits in daily life. This leads to the cause of many kinds of diseases to human. Machine learning, which is an important technique of artificial intelligence, has the ability to learn also predict the diseases automatically. The aim of deep learning in healthcare is its capability to extract the features from big number of datasets. Globally, there is a substantial unmet require to diagnose various diseases effectively. The complexity of the different disease mechanisms and underlying symptoms of the patient population presents massive challenges in developing the early diagnosis tool also effective treatment. Machine learning (ML), an area of artificial Intelligence (AI), enables researchers, physicians, and patients to solve some of these type of issues. Based on different types of research, this review explains how machine learning (ML) is being used to help in the earlier stage identification of numerous diseases. Based on relevant research, this review explains how machine learning (ML) is being used to help in the early identification of numerous diseases of humans.**

Keywords: Machine Learning, Deep Learning, Disease Detection, Pattern matching, image processing, CNN, Internet of Things (IoT), Big data (BD) tools, Disease prediction, symptoms.

I.INTRODUCTION

In medical domains, artificial intelligence (AI) primarily focuses on developing the algorithms and techniques to determine whether a system's behaviour is correct in term of disease diagnosis. Medical diagnosis identifies the disease or conditions that explain an person's symptoms and signs. Typically, diagnostic information is gathered from the patient's history and physical examination of human. The disease occurs when the cells in a healthy person body are impaired due to infection or illness. Breast Cancer diseases, COVID-19, Brain tumor diseases, and Diabetes diseases are the major diseases that are considered as life threatening diseases in day by day. They are likely to be communicable or also non-communicable diseases. Breast Cancer diseases affect the lungs of the human body either in a larger level or can be smaller level. Due to rapid advancements in computer technologies, they becomes easy to predict any type of disease occurring among the common peoples. Artificial Intelligence plays an important role in the area of Computer Science which makes machines mimic the human brain and to act more intelligently.

A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm that can takes in an input image, assign importance (learnable weights and biases) to various aspects or objects in the image, and be able to differentiate one from the other. The main advantages of this project is that we can get the test results immediately at our home with a just few clicks from respective sites.

Medicines and healthcare are some of the most crucial parts of the economy and human life strategy. There is a tremendous amount of change in the world we are living in now and the world that existed a few weeks are back. Everything has turned gruesome and different. In this condition, where everything has turned virtual, the doctors and nurses are putting up maximum efforts to save people's lives even if they have to danger their own also. There are also some remote villages which lack medical facilities to human. Virtual doctors are board-certified doctors who choose to practice online via video and mobile phone appointments, rather than in-person appointments but this is not possible in the case of emergency.

Machines are always considered better than humans as, without any human error, they can perform tasks more correctly and with a consistent level of accuracy. A disease predictor can be called a virtual doctor, which can predict the disease of any patient without any type of human error. Also, in conditions like COVID-19 and EBOLA, a disease predictor can be a blessing as it can identifies a human's disease without any physical contact. Some models of virtual doctors do exist, but they do not comprise the required level of accuracy as all the parameters required are not well being considered. The primary goal was to develop numerous models to done which one of them provides the most accurate predictions outputs.

While ML projects vary in scale and complexity, their general structure be the same. Several rule-

based techniques were drawn from machine learning to recall the development and deployment of the predictive models. Several models were initiated by using various machine learning (ML) algorithms that collected raw data from users and then bifurcated it according to gender, age group, and symptoms. The data-set was then processed in several ML models like Fine, Medium and Coarse Decision trees, Gaussian Naive Bayes, Kernel Naive Bayes, Fine, Medium and also Coarse KNN, Weighted KNN, Subspace KNN, and RUS Boosted trees. According to ML models, that accuracy varied. While processing the data, the input parameters data-sets was supplied to every model, and the diseases was received as an output with correct results.

• At present, when one suffers from particular disease, then the person has to visit to doctor which is time consuming and costly too. Also if the user is out of reach of doctor and hospitals it may be difficult for the user as the disease can not be identified. So, if the above process can be completed using a automated program which can save time as well as money, it could be easier to the patient which can make the process easier .

• This project is design to detect multiple disease such as Brain tumor ,Alzheimer disease, Breast Cancer , Skin Cancer.

• Diabetic Retinopathy (DR) is a common complication of diabetes mellitus, which causes lesions on the retina that effect vision. If it is not detected early, it can lead to blindness.

• ALZHEIMER'S disease (AD) is the leading cause of dementia and poses a significant social and economic challenge. It is responsible for more than half of all cases of dementia.

• Brain tumor can be classified into two types: benign and malignant. Timely and prompt disease detection and treatment plan leads to improved quality of life and increased life expectancy in these patients. One of the most practical and important methods is to use Deep Neural Network (DNN).

• The most frequently occurring cancer among Indian women is breast cancer. There is a chance of fifty percent for fatality in a case as one of two women diagnosed with breast cancer die in the cases of Indian women.

II.MOTIVATION

The purpose of this research is to provide insights to recent and future researchers and practitioners regarding machine-learning-based disease diagnosis (MLBDD) that will aid also enable them to choose the most appropriate and superior machine learning or deep learning methods.

Human disease can affect almost every part of your body. Early detection of Brain tumor, breast cancer, Alzhimer disease and diabetic retinopathy should be beneficial because it enables clinicians to initiate effective treatment of mild disease, preventing loss of body part function and delaying or avoiding progression to organs failure.

III.ALGORITHM

• Convolutional Neural Networks (CNN) –

As in any other neural network, the input of a CNN, in this case an image, is passed through a series of filters in order to obtain a labelled output that can then be classified. The specificity of a CNN lies in its filtering layers, which include at least one convolution layer. These allow it to process more complex pictures than a regular neural network. Whereas the latter is well adapted for simple, well-centred images such as hand-written digits, the use of CNNs in image analysis

ranges from Facebook's automatic tagging algorithms, to object classification and detection, in particular in the field of radiology.

There are Four types of layers in Convolutional Neural Networks:

1) Convolutional Layer: In a typical neural network each input neuron is connected to the next hidden layer. In CNN, only a small region of the input layer neurons connect to the neuron hidden layer.

2) Pooling Layer: The pooling layer is used to reduce the dimensionality of the feature map. There will be multiple activation pooling layers inside the hidden layer of the CNN.

3) Flatten: Flattening is converting the data into a 1-dimensional array for inputting it to the next layer. We flatten the output of the convolutional layers to create a single long feature vector.

4) Fully-Connected layer: Fully Connected Layers form the last few layers in the network. The input to the fully connected layer is the output from the final Pooling or Convolutional Layer, which is flattened and then fed into the fully connected layer.

Four types of disease detections:

- Brain Tumour Detection
- Breast Cancer Detection
- Alzheimer Detection
- Skin Cancer Detection

IV.METHODOLOGY

In our approach, we pre-process the data and select the most important features. Only those features are selected and the rest are eliminated. The model is trained and tested using various models like SVM, Random forest and hybrid neural network model. The efficiency of various algorithms are compared using different performance parameters.

1.**Dataset** – Provide dataset (This means that the data collected should be made uniform and understandable for a machine that doesn't see data the same way as humans do.)

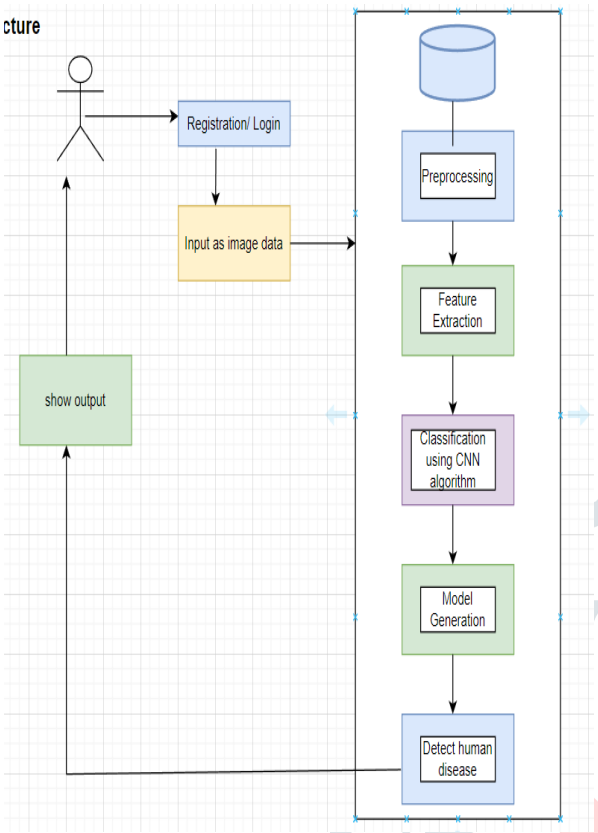
2.**Pre-processing** –A real-world data generally contains noises, missing values, and maybe in an unusable format which cannot be directly used for machine learning models. Data pre-processing is required tasks for cleaning the data and making it suitable for a machine learning model which also increases the accuracy and efficiency of a machine learning model.

3.**Feature Extraction** – Feature Extraction aims to reduce the number of features in a dataset by creating new features from the existing ones (and then discarding the original features). These new reduced set of features should then be able to summarize most of the information contained in the original set of features.

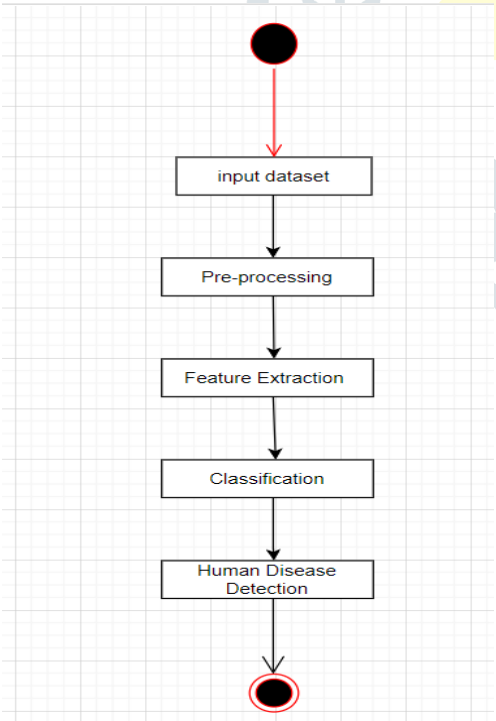
4.**Classification** - The Classification algorithm is a Supervised Learning technique that is used to identify the category of new observations on the basis of training data. In Classification, a program learns from the given dataset or observations and then classifies new observation into a number of classes or groups.

V.SYSTEM DIAGRAMS

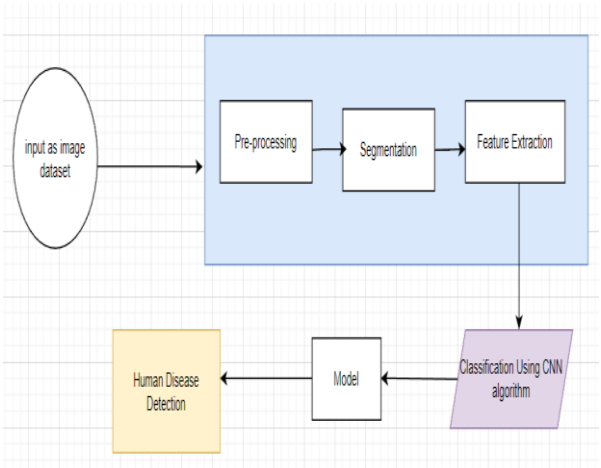
A. System Architecture



B. Activity Diagram



C. Deployment Diagram



VI.CONCLUSION

This healthcure system consists of classification and prediction machine learning algorithms. The machine learning algorithms used are decision tree, random forest and svm algorithms are successfully implemented and they are able to classify the dataset and predict the accuracy of each algorithm by using each disease dataset. The chronicle dataset Heart diseases, Diabetes Mellitus dataset, Liver dataset will used in the phase of implementation. All these disease dataset are of different data size and with different numbers of attributes are used to check the efficiency of machine learning algorithms, as we can see in the comparison table and the average of all algorithms for different dataset then we get that the Random forest generates highest accuracy than in comparison with other. This implemented application helps in medical field as an lot of that is stored in medical care industry, to early diagnosis of disease and to help the patients this application can be helpful.

As we have already seen the applications of data mining and machine learning in medical sector. In this project, a CNN is implemented for detection of Human disease . Although the classifiers worked efficiently in prediction of other diseases

also. In this project, Human Disease is predicted using CNN Algorithm and a comparative study of their performance is done. From the analysis we found that, out of one classifiers. CNN classifier performed better than the other. The rate of detection of human disease is improved.

VII.FUTURE SCOPE

Once a particular form of heart disease is diagnosed, the patient has several types of treatment options. Data extraction is very useful in ML. This is a good data set approach. Therefore, different models need to be combined with more complex models to improve the accuracy of predicting the risk of heart disease in human beings. This system is very intelligent as more data is added to the database system. Many improvements there can be made to improve scalability and accuracy results. There using different interpretation methods, different types of decision trees increase information and profitability, meaning multiple classification voting methods.

As illustrated before the system can be used as a clinical assistant for any type of clinicians. The disease prediction through the risk factors can be hosted online and hence any internet users can access the system through a web browser and understand good as well the risk of heart disease. The proposed model requires an efficient processor with good memory configuration to implement it in real time efficiently. The proposed model has wide area of application like an grid computing, cloud computing, robotic modeling, etc.

REFERENCES

1. M. Karuna and A. Joshi, Automatic detection and severity analysis of brain tumors using gui in matlab, International Journal of Research in Engineering and Technology, 10, pp. 586-594, 2013.
2. KS. Aboody, A. Brown, et al, Neural stem cells display extensive tropism for pathology in adult brain Evidence from intracranial gliomas,Proceedings of the National Academy of Sciences, 97 (23), pp. 1284612851, 2000
3. M. Karuna and A. Joshi, Automatic detection and severity analysis of brain tumors using gui in matlab, International Journal of Research in Engineering and Technology, 10, pp. 586-594, 2013.
4. KS. Aboody, A. Brown, et al, Neural stem cells display extensive tropism for pathology in adult brain Evidence from intracranial gliomas, Proceedings of the National Academy of Sciences, 97 (23), pp. 1284612851, 2000
5. L. Tribuna et al.A study to understand the acceptance of DICOM Structured Reports on Breast Imaging Proc. Comput. Sci.(2017)
6. J.M. Silva et al.Controlled searching in reversibly de-identified medical imaging archives J. Biomed. Inf.(2018)
7. C. Huang et al., “Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China,” Lancet, vol. 395, no. 10223, from pp. 497–506, 2020.
8. R. H. Abiyev and M. K. S. Ma’aitah, “Deep Convolutional Neural Networks for Chest Diseases Detection,” J. Healthc. Eng., vol. 2018, from pp. 1–11, Aug. 2018, doi: 10.1155/2018/4168538.
9. R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, “Convolutional neural networks:

an overview and application in radiology," Insights Imaging, vol. 9, no. 4, pp. 611–629, Aug. 2018, from doi: 10.1007/s13244018-0639-9.

10. Y.-H. Chan, Y.-Z. Zeng, H.-C. Wu, M.-C. Wu, and H.-M. Sun, "Effective Pneumothorax Detection for Chest X-Ray Images Using Local Binary Pattern and Support Vector Machine," J. Healthc. Eng., vol. 2018, pp. 1–11, 2018, from doi: 10.1155/2018/2908517.

11. Kumar, R.R.; Polepaka, S. Performance Comparison of Random Forest Classifier and Convolution Neural Network in Predicting Heart Diseases. In ICCII 2018, Proceedings of the Third International Conference on Computational Intelligence and Informatics; Springer: Singapore, 2020; from pp. 683–691.