WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Disease Prediction Using Machine Learning and Deep Learning Techniques

Aftab Attar [(1)], Akash Rathod [(2)], Anirudh V [(3)], MadanBeleri [(4)]

Department of Information Science and Engineering, Dayananda Sagar College of Engineering
Asst. Professor, Prathima Mabel J, Department of Information Science and Engineering, Dayananda Sagar College of Engineerin

ABSTRACT

Diseases are the main reason for a large number of deaths in the world over the last few decades. It has emerged as one of the most formidable life-threats to mankind, not only in India but over the entire world. This creates a need for a reliable and accurate system to diagnose such diseases in time for proper treatment. Machine Learning techniques have been applied to various medical datasets to automate the analysis of large, complex data.In this project, an efficient, automated disease diagnosis system will be designed using machine learning and deep learning models. Various critical diseases like malaria, pneumonia, cancer etc. have been chosen during the course of the project. The data will be entered into the system and analysis will then be performed using a pre-trained ML/DL model which is trained with corresponding datasets. Finally, the image provided is classified by Sequential CNN into either a disease-afflicted or normal image and the prediction results are displayed to the user.According to research, there are 40% of people who ignore primary symptoms first which leads to complications later. The main reason for ignorance is their busy schedules. The main motive of our project is to develop a system where a user can sit at their convenient place and can monitor their health.This system helps people in getting their prediction done with good accuracy.

Keywords: Disease prediction, Sequential CNN, Image Classification.

INTRODUCTION

Disease Prediction using Machine Learning and Deep learning is a system that predicts the disease based on the information provided by the user. It also predicts the disease of the patient or based on the information he/she enters into the system and provides accurate results based on it. The Health Industry plays a major role in curing the diseases of the patients and our system in turn aids the health industry through disease identification and prediction. Similar models were previously built by many other people but were pertaining only to isolated diseases. Here, our intention is to make it different and beneficial for the users by incorporating various prevalent diseases into a single, effective system. This system is built with the help of Python Programming language and its libraries. We will predict the disease using real-time datasets.Therefore, disease prediction systems that use machine learning algorithms assist in cases which require a second opinion to get accurate results.

In the competitive environment of economic development, mankind has become so workaholic that people are not concerned about their health. According to research, there are 40% of people who ignore primary symptoms first which leads to complications later. The main reason for ignorance is their busy schedules. People have become so hard-pressed for time that they have no bandwidth to make an appointment and consult the doctor.

LITERATURE SURVEY

1. **"Heart Disease Prediction Using Machine Learning Techniques "by Dr Poonam Ghuli** In this paper theyhave proposed a system which predicts the chances of Heart Disease and classifies patient's risk level by implementing different data mining techniques such as Naive Bayes, Decision Tree, Logistic Regression and Random Forest. The trial results verify that the Random Forest algorithm has achieved the highest accuracy of 90.16% compared to other ML algorithms implemented.
2. **"Impact of machine learning and feature selection on type 2 diabetes risk prediction "by Riihimaa P.**
   This survey summarizes the state of the art for type 2 diabetes mellitus (T2DM) prediction and compares the prediction accuracies obtained by conventional statistical regression and machine learning methods, including deep learning. The impact of feature selection and inclusion of clinical and genomic data on T2DM risk prediction accuracy is also reviewed.
3. **"Cancer Prediction Using Machine Learning", Journal of Emerging Technologies and Innovative Research (2020) by Ramik Rawal**
   This model is designed to counter the lack of robust prognosis models which results in difficulty for doctors to prepare an effective treatment plan that may prolong patient survival time. Four algorithms SVM, Logistic Regression, Random Forest and KNN which predict the breast cancer outcome have been compared in the paper using different datasets. Aim of research categorizes in three domains, 1st domain is prediction of cancer before diagnosis, 2nd domain is prediction of diagnosis and treatment and the 3rd domain focuses on outcome during treatment.
4. **"Deep Machine Learning Model Trade-Offs for Malaria Elimination in Resource-Constrained Locations" Eze, P.U.; Asogwa, C.O**
   The goal is to maximize malaria detection accuracy while reducing computing resources and energy consumption. Based on our experimental results using a blood smear malaria test data set, the quantized versions of Basic Convolutional Neural Network (B-CNN) and MobileNetV2 have better malaria detection performance, lower memory usage (2MB 8-bit quantized model) and shorter inference time (33–95 microseconds) than VGG-19 fine-tuned and quantized models.

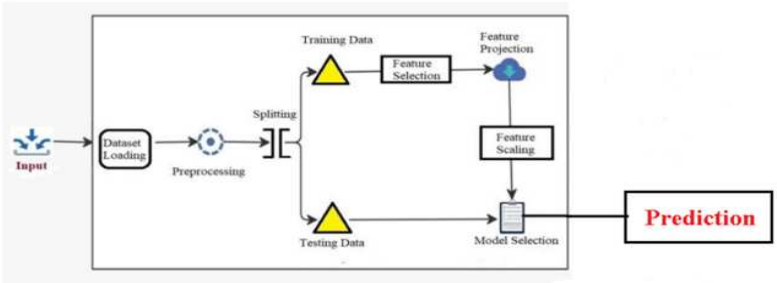
5. "Pneumonia Detection Using Deep Learning Based on Convolutional Neural Network" (2021) by Luka Račić, Tomo Popović

This paper describes the use of machine learning algorithms to process chest X-ray images in order to support the decision-making process in determining the correct diagnosis. Specifically, the research is focused on the use of deep learning algorithms based on convolutional neural networks in order to build a processing model. This model has the task to help with detecting whether a chest X-ray shows changes consistent with pneumonia or not.

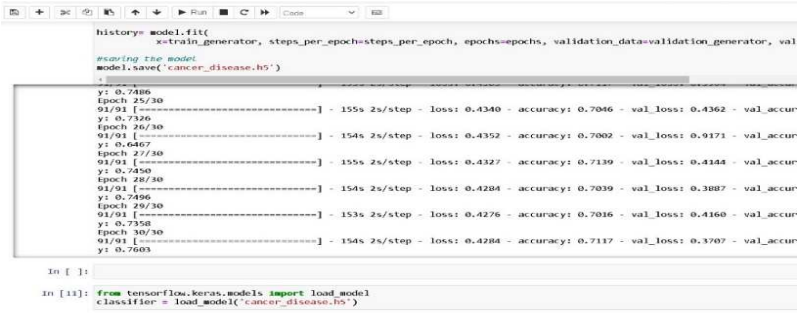
PROPOSED SYSTEM

The system that we have developed is cost effective and reliable as the person can try to predict the disease easily.Keeping in mind some of the drawbacks of the existing system, we have proposed a system that performs significantly better in terms of usage of data and computational resources.Our proposed system uses CNN algorithm to train and evaluate the obtained model under available datasets.The system mainly focuses on predicting some prevalent diseases rather than one disease.This system not only predicts but is also aimed at predicting with better accuracy.The system is built using a single CNN model for disease prediction of all the chosen diseases.Separate two classes are built in for prediction of each disease.The CNN used is sequential CNN which includes 5 convolutional layers and 3 fully connected layers with batch normalization, activation,padding and max pooling functions.

The system is built using python language along with its suitable modules such as TensorFlow and Keras.The actual architecture of the system includes preprocessing of datasets collected followed by splitting of data into training and testing and then feature selection,feature projection and feature scaling is done in order to obtain best model which is then fed with tested data.

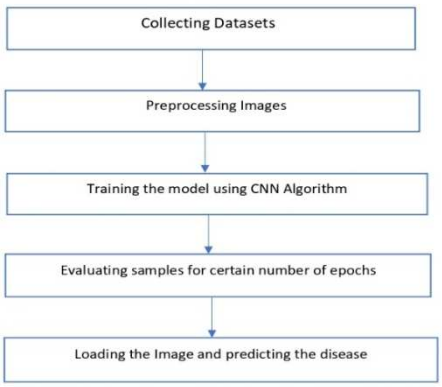


DATASET



The main effort put forth was in collecting some real time datasets from known family doctors and medical practitioners which helped us a lot in training the model with them.The image datasets for each disease are collected and stored into separate folders as positive and negative images. 5842 images of all diseases combined were collected and used for training the model.This includes both X-ray and CT scan images.

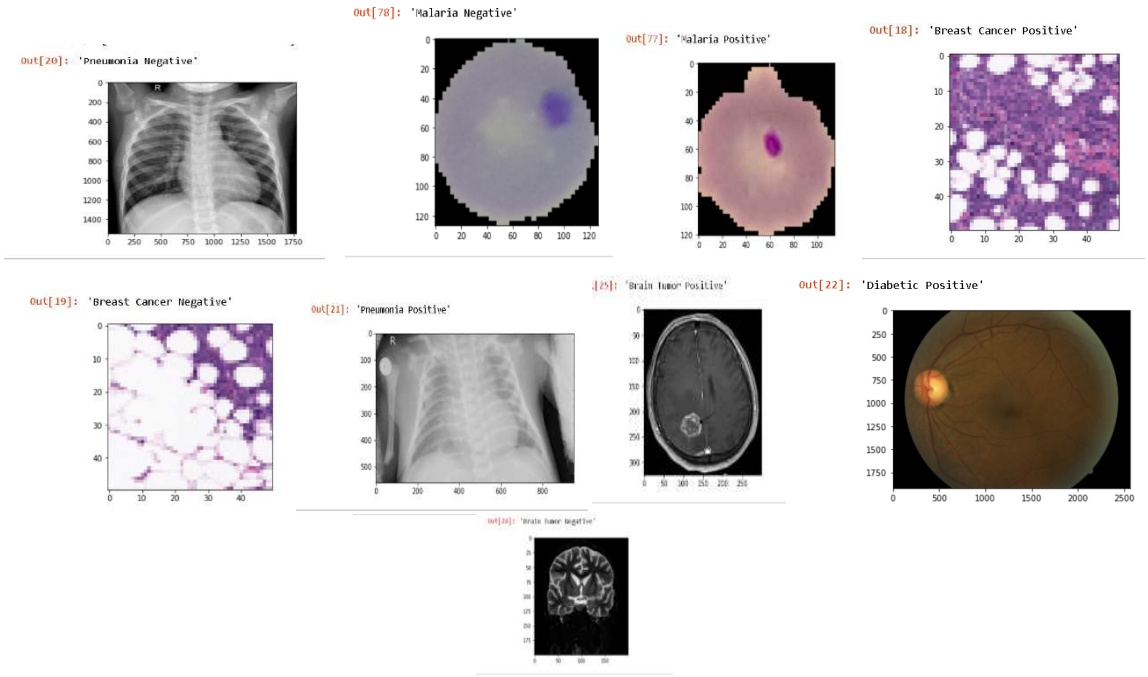
SEQUENCE DIAGRAM



Coming to the sequence diagram of our project, the first step was the collection of real time datasets along with some available on the internet, followed by preprocessing of the images into system understandable format. Next step included the building and training of a model created using CNN algorithm.The model was further trained using the tested data for a certain fixed number of epochs. Later on the images were loaded into the system for further prediction.

RESULTS

The images shown below are the results of our predictionson 5 prevalent diseases. We have obtained an accuracy of over 75% while training the model for 30 valid number of epochs. Since, Sequential CNN has been used, the maximum epoch limit for training the model is 60 post which the training stops.



CONCLUSION

The fast-paced lifestyle of today has had a profound effect on the well-being of people worldwide and the number of diseases and related deaths continues to grow globally. Therefore, disease prediction systems that use machine learning algorithms assist in cases which require a second opinion to get accurate results.Machine Learning and Deep Learning are possible successful ways to provide promising data driven solutions to help humanity in handling diseases. In this project, we have explored the ML methods to diagnose diseases and also summarized the available datasets, tools and performance. A few challenges were also faced in the process.

FUTURE SCOPE

The efficacy of Machine Learning models and its graphical characteristics can be enhanced in order to predict more diseases.This will help radiologists and practitioners in early prediction and diagnosis.Similarly,in order to find the treatment for different diseases future research work on ML and DL is required.Also, further research on the techniques can be done and a system based on different specifications can be built to predict and treat the disease.

REFERENCES

1. Dr Poonam Ghuli et al.,” Heart Disease Prediction Using Machine Learning Techniques”, International Journal of Engineering Research & Technology, Vol. 9 Issue 04, April-2020.
2. Rihiimaa P et al.,” Impact of machine learning and feature selection on type 2 diabetes risk prediction”, Journal of Medical Artificial Intelligence, 2020.
3. Ramik Rawal et al., “Cancer Prediction Using Machine Learning”, Journal of Emerging Technologies and Innovative Research (2020), Vol. 7, Issue 5.
4. Eze, P.U.; Asogwa, C.O Deep Machine Learning Model Trade-Offs for Malaria Elimination in Resource-Constrained Locations. Bioengineering 2021.
5. Račić, Luka & Popovic et al. “Pneumonia Detection Using Deep Learning Based on Convolutional Neural Network” (2021)25th International Conference on Information Technology, Montenegro.
6. D. Ravì et al., "Deep Learning for Health Informatics,", IEEE Journal of Biomedical and Health Informatics, 21(1), 2017, pp. 4-21.
7. Suresha D, Jagdish N, Shrisha H &Kaushik KS, “Detectionof Brain Tumor Using Image Processing”,2020.
8. Parsa Moradi, Amirhosein Khas Ahmadi,” Discovering Informative Regions in Iris Images to Predict diabetes”,2018.