International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

ISCHEMIC STROKE PREDICTION USING MACHINE LEARNING ALGORITHMS

[1]Mr.G.Prabhakar Raju, [2]I.Malleshwari, [3]M.Vennala, [4]P.Harsha Vardhan Reddy
[1]Assistant Professor,Dept of Computer Science and Engineering,
Anurag Group of Institutions TS,India.
[2,3,4]Students, Computer Science and Engineering,Anurag Group of Institutions,TS,India.
prabhakarrajucse@cvsr.ac.in,19h61a05d9@cvsr.ac.in.19h61a05e9@cvsr.ac.in ,
19h61a05f4@cvsr.ac.in

ABSTRACT:

An ischemic stroke is a medical condition in which blood supply to the brain is interrupted or reduced, preventing brain tissue from getting oxygen and nutrients that results in cell death. It is now-a-days leading to cause of death all over the world. Several risk factors believe to be related to the cause of stroke has been found by inspecting the affected individuals. Using these risk factors, a number of works have been carried out for predicting the stroke diseases. Early recognition of the various warning signs of a stroke can help reduce the severity of the stroke. Machine Learning delivers an accurate and quick prediction outcome. It's applications in health care are growing, nonetheless there is a greater need for further investigation in some research fields. We believe that machine learning algorithms can help in better understanding of diseases and can be a good healthcare companion. For this work the textual dataset of various patients which contains various medical factors is collected. The dataset is processed by identifying and handling the missing values. The machine learning algorithms such as Random Forest, Decision Tree classifier and SVM are trained on the dataset. Then the accuracy of the algorithms is found and the algorithm that provides the best accuracy for our dataset is been chosen. This helps the patients in assessing the possibility of brain stroke and make sure to take proper health care.

Keywords: strock, ML, SVM, MD, NHIS,

I. INTRODUCTION

According to the World Stroke Organization, 13 million people get a stroke each year, and approximately 5.5 million people will die as a result. It is the leading cause of death and disability worldwide, and that is why its imprint is serious in all aspects of life. Stroke not only affects the patient but also affects the patient's social environment, family, and workplace. It is the fourth largest cause of death in India. Because of developments in medical technology, Machine Learning can now predict the occurrence of a stroke. Machine Learning algorithms are useful for making accurate predictions and analysing data. Most of the previous stroke research has focused on predicting heart attacks. Little or no attention has been paid to brain stroke. The algorithms that exist in machine learning are constructive to make

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

accurate predictions and provide correct analysis. The work done so far on the topic of stroke mainly includes work on heart rate prediction. Little research has been done on stroke. The project is based on using machine learning to predict the occurrence of stroke.

Machine Learning is one of the efficient technologies for the testing, which is based on training and testing. It is the branch of Artificial Intelligence (AI) which is one of broad area of learning where machines emulating human abilities, machine learning is a specific branch of AI. On the other hand, machines learning systems are trained to learn how to process and make use of data hence the combination of both technology is also called as Machine Intelligence. As the definition of machine learning, it learns from the natural phenomenon, natural things so in this project we use the biological parameter as testing data such as hypertension, heart disease, sex, age, etc.A dataset with various physiological characteristics has been selected as. These features will be analysed later and used for final predictions. The dataset is first cleaned up and prepared to understand the machine learning model. This step is called data preprocessing. To do this, the zero values in the data record are checked and these are entered. After pre-processing the data, the dataset is split into training and test data. This new data is then used to build a model using various classification algorithms. The accuracy of all these algorithms is calculated and compared to obtain the best trained model for the prediction. The limitation of this model is that it is trained with matter information instead of time period brain picture.

II. LITERATURE SURVEY

1. Medical software user interfaces, stroke MD application design (IEEE).

AUTHORS: Elena Zamsa

The article presents the design of an application interface for associated medical data visualization and management for neurologists in a stroke clustering and prediction system called Stroke MD. The goal of the system is to facilitate efficient visual data introduction and knowledge extraction based on a predictive model implementation. Aspects such as quality of the visualized data, type of the viewed data (risk factor groups, clustered groups, data alerts, conclusions etc.) and efficient usersoftware interaction directly influenced decisions about system design, in particular the design of the doctor's/patient's user interface sides, input methods and usersoftware interaction techniques, ending in to a hybrid setup of interface.

Development of an Algorithm for Stroke Prediction: A National Health Insurance Database Study.

AUTHORS: Min SN, Park SJ, Kim DJ, Subramaniyam M, Lee KS

In this research, this paper aimed to derive a model equation for developing a stroke pre- diagnosis algorithm with the potentially modifiable risk factors. It used logistic

regression for model derivation, together with data from the database of the Korea National Health Insurance Service (NHIS). It reviewed the NHIS records of 500,000 enrolees. For the regression analysis, data regarding 367 stroke patients were selected. The control group consisted of 500 patients followed up for 2 consecutive years and with no history of stroke. It developed a logistic regression model based on information regarding several well-known modifiable risk factors. The developed model could correctly discriminate between normal subjects and stroke patients in 65% of cases. The model developed in the present study can be applied in the clinical setting to estimate the probability of stroke in a year and thus improve the stroke prevention strategies in high-risk patients. The approach used to develop the stroke prevention algorithm can be applied for developing similar models for the prediagnosis of other diseases.

Probability of Stroke: A Risk Profile from the Framingham Study.

AUTHORS: Philip A. Wolf, MD; Ralph B. D'Agostino, PhD, Albert J. Belanger,
MA; and William B. Kannel, MD.

In this paper, A health risk appraisal function has been developed for the prediction of stroke using the Framingham Study cohort. A health risk appraisal function has been developed for the prediction of stroke using the Framingham Study cohort. The stroke risk factors included in the profile are age, systolic blood pressure, the use of antihypertensive therapy, diabetes mellitus, cigarette smoking, prior cardiovascular disease (coronary heart disease, cardiac failure, or intermittent claudication), atrial fibrillation, and left ventricular hypertrophy by electrocardiogram. Based on 472 stroke events occurring during 10 years' follow-up from biennial examinations 9 and 14, stroke probabilities were computed using the Cox proportional hazards model for each sex based on a point system. Based on the risk factors in the profile, which can be readily determined on routine physical examination in a physician's office, stroke risk can be estimated. The information that one's risk of stroke is several times higher than average may provide the impetus for risk factor modification. It may also help to identify persons at substantially increased stroke risk resulting from borderline levels of multiple risk factors such as those with mild or borderline hypertension and facilitate multifactorial risk factor modification.

III.EXISTING SYSTEM

Stroke patients commonly use Ayurveda therapies, but there are no published data regarding the efficacy or safety of these therapies. Ayurvedic therapies lacks of accuracy to finding out the stroke disease. This accuracy can be improved by using machine learning techniques. People cannot predict a stroke or a risk of having stroke unless they suffer with it. Very few systems use the available clinical data for prediction purposes and even if they do, they are restricted by the

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

large number of association rules that apply. Diagnosis of the condition solely depends upon the Doctor's intuition and patient's records.

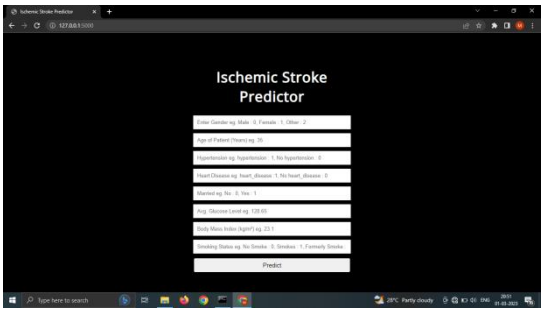
IV.PROPOSED SYSTEM

To deal with a problem there is an essential need of prediction system for awareness about diseases. Machine learning is the branch of Artificial Intelligence (AI), it provides prestigious support in predicting any kind of event which takes training from natural events. In proposed system, for prediction machine learning techniques are used such as Logistic Regression, Random Forest and SVM. Training the models and finding the accuracy, to choose an algorithm that provides the best accuracy for the dataset. This technology can estimate the likelihood of a patient developing an illness by looking at their medical profiles, including age, blood pressure, sugar levels, and more. Classification algorithms are used to predict disease when there are many variables.

WORKING

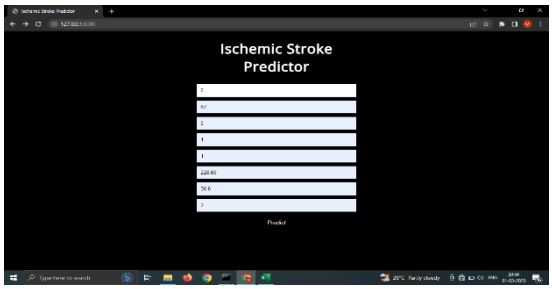
The proposed system helps us to analyse the best possible way to take inputs from the user in the GUI implementation part of the project shown in the Figure and with the data that is provided to the GUI for stroke risk prediction, the SVM model, which was trained with the dataset, is used and the new data provided by the user is tested against the trained model.

HOME PAGE:



STROKE:

In this GUI, a HTML form is presented to the user wherein the user after entering the required attributes, will be predicting whether the user is at risk of suffering from a brain stroke or not.



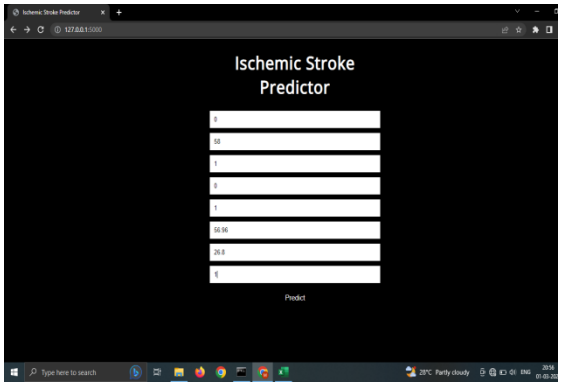
If the user is predicted with no stroke i.e, the user is identified to be not at a risk of suffering from a stroke the below screen will be displayed:

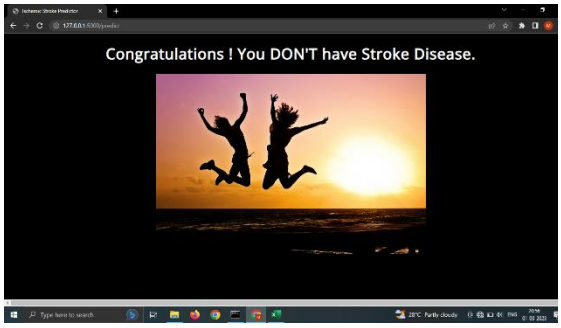
NO STROKE:

In this GUI, a HTML form is presented to the user wherein the user after entering the required attributes, will be predicting whether the user is at risk of suffering from a brain stroke or not. If the user is predicted with stroke

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

i.e, the user is identified to be at a risk of suffering from a stroke the below screen will be displayed:





V.CONCLUSION

The proposed system helps to predict brain strokes in a cost effective and efficient way by taking few inputs from the user side and predicting accurate results with the help of trained Machine Learning algorithms. The project is focused on creating Machine Learning model on assessing three algorithms (Decision tree, SVM, Random Forest) and deploy the one among three to the flask application in identifying stroke-prone patients. Thus, the Brain Stroke Prediction system has been implemented using the given three Machine Learning algorithm given a highest accuracy of 94.30%. The system is therefore designed providing simple yet efficient User Interface design with an empathetic approach towards their users and patients. It helps to predict the stroke risk using prediction model and provide personalized warning and the lifestyle correction message. Finally, various preventative steps such as quitting smoking, avoiding alcohol, and other factors are recommended to reduce the risk of having a stroke.

VI. FUTURE ENHANCEMENT

The system has a potential for future scope which could lead to better results a better user experience. This will help the user to save their valuable time and will help them to take appropriate measures based on the results provided. The future scope for the implemented system can be Increasing the accuracy of the model. To improve the results,a dataset with sufficient features and increases in quantity must be obtained.Further research must be conducted in enhancing the existing machine learning techniques to work in real time and develop an efficient model.Also, the model developed must be tested on data with different volumes to test its performance.

VI. BIBLIOGRAPHY

1. Manisha Sirsat, Eduardo Ferme, Joana Camara, “Machine Learning for Brain
Stroke: A Review,” Journal of stroke and cerebrovascular diseases: the official journal of National Stroke Association (JSTROKECEREBROVASDIS), 2020.
2. Harish Kamal, Victor Lopez, Sunil A. Sheth, “Machine Learning in Acute

Ischemic Stroke Neuroimaging," Frontiers in Neurology (FNEUR), 2018.

3. Chuloh Kim, Vivienne Zhu, Jihad Obeid and Leslie Lenert, "Natural language processing and machine learning algorithm to identify brain MRI reports with acute ischemic stroke," Public Library of Science One (PONE), 2019.
4. R. P. Lakshmi, M. S. Babu and V. Vijayalakshmi, "Voxel based lesion segmentation through SVM classifier for effective brain stroke detection," International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), 2017.
5. J. Yu et al., "Semantic Analysis of NIH Stroke Scale using Machine Learning Techniques," International Conference on Platform Technology and Service (PlatCon), 2019,
6. Gangavarapu Sailasya and Gorli L Aruna Kumari, "Analyzing the Performance of Stroke Prediction using ML Classification Algorithms," International Journal of Advanced Computer Science and Applications (IJACSA), 2021.
7. "Stroke Prediction Dataset". Kaggle.Com, 2021, https://www.kaggle.com/fedesoriano/stroke prediction-dataset. Accessed 6 Oct 2021