International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 10 Oct 2021, pp: 813-819 www.ijaem.net ISSN: 2395-5252

Brain Stroke Prediction Bachelor Of Technology In Computer Science Engineering

Bhavyanshu Ghariya, Satwik Rai, YashSawant, Amit Sharma, Mr. Sachin Jain
(ASSISTANT PROFESSOR)

Submitted: 10-10-2021 Revised: 19-10-2021 Accepted: 22-10-2021

ABSTRACT —A Stroke is a health condition that causes damage by tearing the blood vessels in the brain. It can also occur when there is a halt in the blood flow and other nutrients to the brain. According to the World Health Organization (WHO), stroke is the leading cause of death and disability globally. Most of the work has been carried out on the prediction of heart stroke but very few works show the risk of a brain stroke. With this thought, various machine learning models are built to predict the possibility of stroke in the brain. This paper has taken various physiological factors and used machine learning algorithms like Logistic Regression, Decision Tree Classification, Random Forest Classification, K-Nearest Neighbors, Support Vector Machine and Naïve Bayes Classification to train five different models for accurate prediction. The algorithm that best performed this task is Logistic Regression that gave an accuracy of approximately 82%

I.INTRODUCTION

Stroke is the fifth-leading cause of death in the worldwide. Stroke is a non-communicable infection that is liable for around 11% of total deaths. It is the fourth significant reason for death in India. With the advancement of technology in the medical field, predicting the occurrence of a stroke can be made using Machine Learning. The algorithms present in Machine Learning are constructive in making an accurate prediction and give correct analysis. This paper is based on predicting the occurrence of a brain stroke using Machine Learning. The key components of the approaches used and results obtained are that among the five different classification algorithms used Naïve Bayes has best performed obtaining a higher accuracy metric. The limitation with this model is that it is being trained on textual data and not on real time brain images. The paper shows the implementation of six Machine Learning classification algorithms. This paper can be further extended to implementing all the current machine learning algorithms. A dataset is chosen from Kaggle with various physiological traits as its attributes to proceed with this task.

These traits are later analyzed and used for the final prediction. The dataset is initially cleaned and made ready for the machine learning model to understand. This step is called Data Preprocessing. For this, the dataset is checked for null values and fill them. Then Label encoding is performed to convert string values into integers followed by one-hot encoding, if necessary. After Data Preprocessing, the dataset is split into train and test data. A model is then built using this new data using various Classification Algorithms. Accuracy is calculated for all these algorithms and compared to get the best-trained model for prediction. After training the model and calculating the accuracy, an HTML page and a Flask application are developed. The web application is for the user to enter the values for prediction. The flask application is a framework that connects the trained model and the web application. After proper analysis, the paper concludes which algorithm is most appropriate for the prediction of stroke

II.PROBLEM STATEMENT

- Stroke is the second leading cause of death worldwide and remains an important health burden both for the individuals and for the national healthcare systems. Potentially modifiable risk factors for stroke include hypertension, cardiac disease, diabetes, and

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 10 Oct 2021, pp: 813-819 www.ijaem.net ISSN: 2395-5252

dysregulation of glucose metabolism, atrial fibrillation, and lifestyle factors.

- Therefore, the goal of our project is to apply principles of machine learning over large existing data sets to effectively predict the stroke based on potentially modifiable risk factors. Then it intended to develop the application to provide a personalized warning on the basis of each user's level of stroke risk and a lifestyle correction message about the stroke risk factors.

III.PROPOSED SYSTEM

The proposed strategy focuses on a novel machine learning procedures for Stroke prediction, thus overcoming the existing problem. Different machine learning methods may not perform equally on the same feature set. Therefore,optimalfeaturesetsfor each machine learning methods were defined systematically. By utilizing multiple machine learning algorithms the model can be used in order to increase the performance and accuracy. jupyter was used to develop the proposed system. The proposed model describes the approach taken to develop the proposed solution which entails pre-processing, feature reduction and the final classification. The collected dataset is given as the input which is pre-processed to remove unwanted rows and columns to produce modified dataset. The data modified is analyzed and is compared to people with stroke and without stroke and the dataset is split into train and test data. The test data is used to get the predictions to produce accuracy report.

IV.ADVANTAGES OF PROPOSEDSYSTEM

- When comparing with existing system user can get more accurate output of prediction.
- The module used most suitable machine learning algorithms for improvised learning and predict the most accurate outcome.
- Easy to Interface
- Flexible

V.SYSTEM ARCHITECTURE

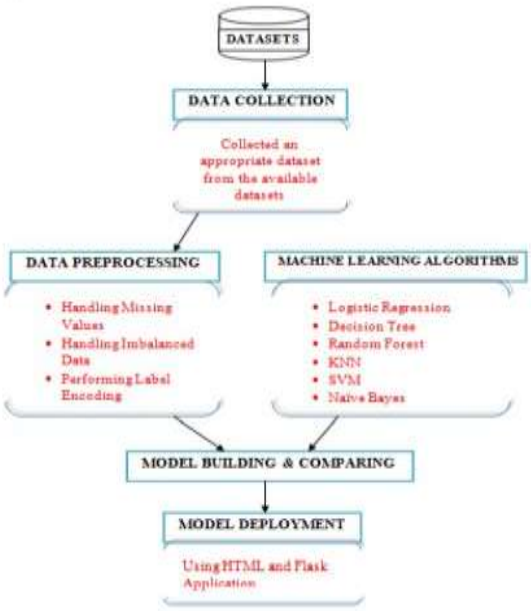

Fig-1: Overall System Architecture

- Dataset will be collected and dat will be cleaned and inconsistency will be removed.
- The most suitable algorithm will be selected from five machine learning algorithms and will be applied.
- The machine will be trained from the cleaned data as 80% data will use for training and 20% for testing.
- The pycharm software will be used for creating the localhost on the computer.
- After model building will be done. The user interface will be created in html.
- The machine learning algorithms will be pickle from the jupyter.
- The data of patient will be entered on application.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 10 Oct 2021, pp: 813-819 www.ijaem.net ISSN: 2395-5252

- The predicted output from given data will be shown by interface after machine learning algorithms will be applied.

VI.DESIGN

DATAFLOW DIAGRAM

The DFD is also known as bubble chart. It is a simple graphical Formalism that can be used to represent a system in terms of the input data to the system, various Processing carried out on these data, and the output data is generated by the system. It maps out the flow of information for any process or system, how data is processed in terms of inputs and outputs. It uses defined symbols like rectangles, circles and arrows to show data inputs, outputs, storage points and the routes between each destination. They can be used to analyse an existing system or model of a new one. A DFD can often visually "say" things that would be hard to explain in words and they work for both technical and non-technical. There are four components in DFD:

1. External Entity
2. Process
3. Data Flow
4. data Store

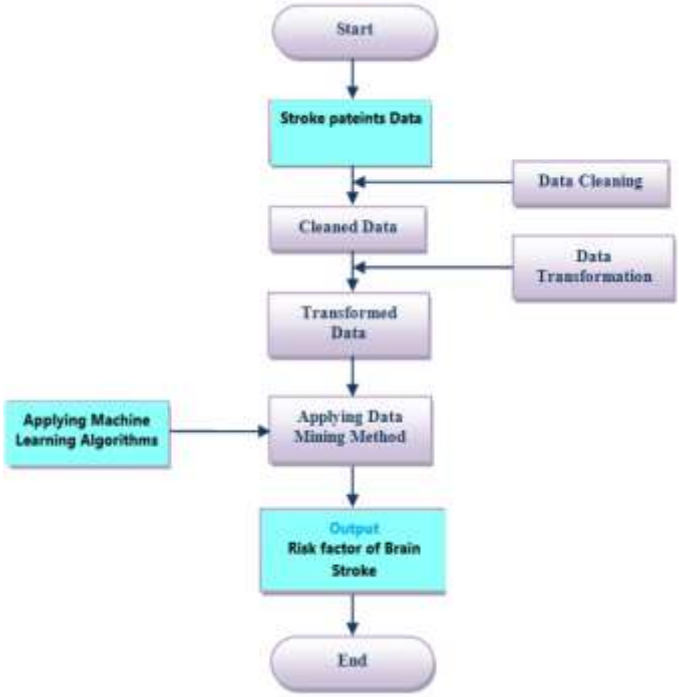

Fig-2:Dataflow Diagram for Brain Stroke prediction

USECASE DIAGRAM

Use Case during requirement elicitation and analysis to represent the functionality of the system. Use case describes a function by the system that yields a visible result for an actor. The identification of actors and use cases result in the definitions of the boundary of the system i.e., differentiating the tasks accomplished by the system and the tasks accomplished by its environment. The actors are outside the boundary of the system, whereas the use cases are inside the boundary of the system. Use case describes the behavior of the system as seen from the actor's point of view. It describes the function provided by the system as a set of events that yield a visible result for the actor. Purpose of Use Case Diagrams The purpose of use case diagram is to capture the dynamic aspect of a system. However, this definition is too generic to describe the purpose, as other four diagrams (activity, sequence, collaboration, and State chart) also have the same purpose. We will look into some specific purpose, which will distinguish it from other four diagrams.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 10 Oct 2021, pp: 813-819 www.ijaem.net ISSN: 2395-5252

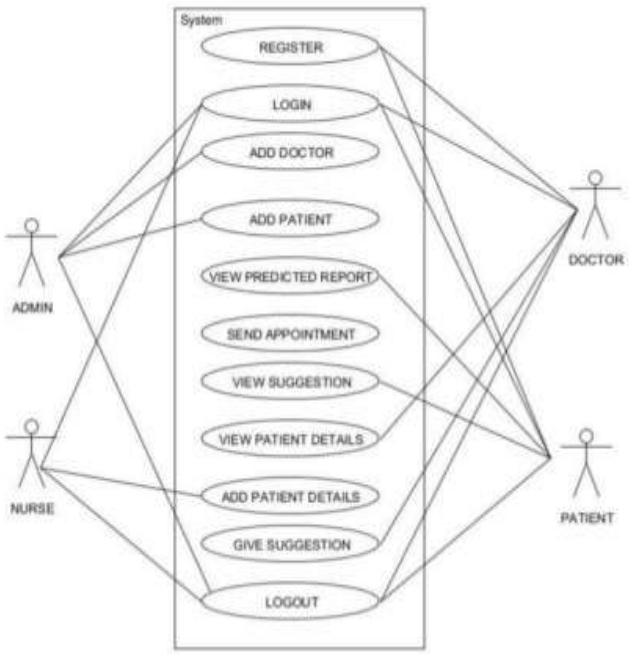

Fig 3: Use case diagram of brain stroke prediction Systemd

Table-1: Usecase Scenario for Brain stroke prediction system

Use Case Name	Brain Stroke Prediction Using Machine Learning
Participating actors	User, System
Flow of the Events	Start the System(U) Entering the patients data(U) Analyzing the data(S) Predicting the output(S)
Entry condition	Run the code, give the patients data
Exit condition	Displaying the result
Quality requirement	Right data values, functioning of system

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 10 Oct 2021, pp: 813-819 www.ijaem.net ISSN: 2395-5252

- CLASS DIAGRAM

Class diagrams model class structure and contents using design elements such as classes, packages and objects. Class diagram describe the different perspective when designing a system conceptual,Specification and implementation. Classes are composed of three things : name,attribute and operations. Class diagram also display relationships such as containment, inheritance, association, etc. The association relationship is the most common relationship in class diagram. The association shows the relationship between instance of classes.

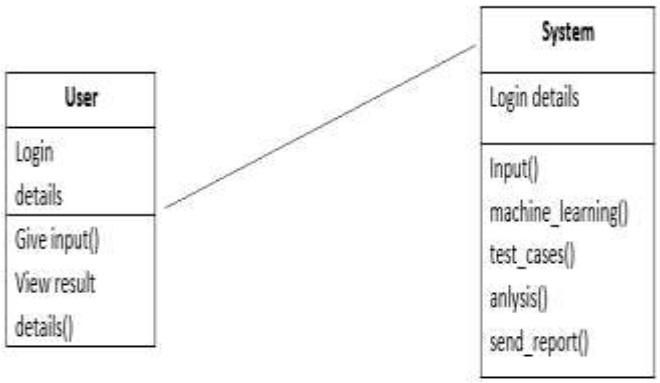

Fig-4: Class diagram of Brain Stroke prediction system

- **SEQUENCE DIAGRAM**

Sequence diagram displays the time sequence of the objects participating in the interaction. This consists of the vertical dimension (time) and horizontal dimension (different objects). Objects: Object can be viewed as an entity at a particular point in time with specific value and as a holder of identity. A sequence diagram shows object interactions arranged in time sequence. It depicts the objects and classes involved in the scenario and the sequence of messages exchanged between the objects needed to carry out the functionality of the scenario. Sequence diagrams are typically associated with use case realizations in the Logical View of the system under development. Sequence diagrams are sometimes called event diagrams or event scenarios. A sequence diagram shows, as parallel vertical lines (lifelines), different processes or objects that live simultaneously, and, as horizontal arrows, the messages exchanged between them, in the order in which they occur. This allows the specification of simple runtime scenarios in a graphical manner.

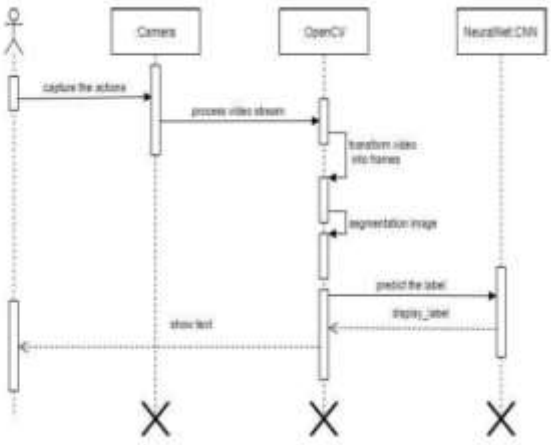

Fig. 5 :- Sequence diagram of brain stroke prediction system

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 10 Oct 2021, pp: 813-819 www.ijaem.net ISSN: 2395-5252

- **STATE CHART**

A state chart diagram describes a state machine which shows the behavior of classes. It shows the actual changes in state not processes or commands that create those changes and is the dynamic behavior of objects over time by model1ing the life cycle of objects of each class. It describes how an object is changing from one state to another state. There are mainly two states in State Chart Diagram: 1. Initial State 2. Final-State. Some of the components of State Chart Diagram are:

State: It is a condition or situation in life cycle of an object during which it's satisfies same condition or performs some activity or waits for some event.

Transition: It is a relationship between two states indicating that object in first state performs some actions and enters into the next state or event.

Event: An event is specification of significant occurrence that has a location in time and space

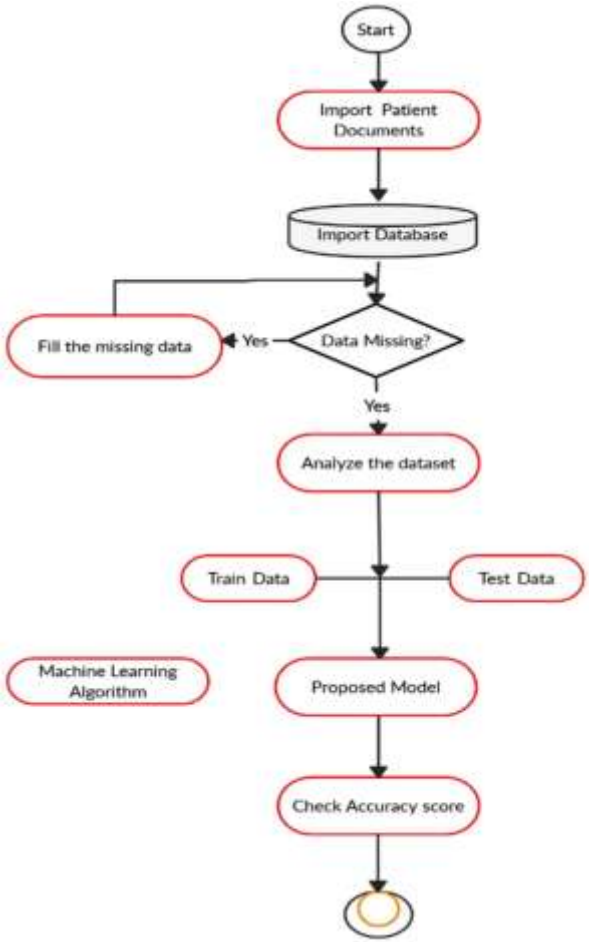

Fig-6: State Chart diagram of brain stroke prediction system

VII. EXPERIMENTAL ANALYSIS AND RESULTS

- SYSTEM CONFIGURATION
- Software requirements Operating System : Windows, Mac, Linux.
- SDK: Pycharm , Jupyter, pandas, Numpy
- Hardware Requirements The Hardware Interfaces Required are:
- Ram: Minimum 8GB or higher
- GPU: 4GB dedicated
- Processor: Intel Pentium 4 or higher
- HDD: 10GB or higher
- Monitor: 15" or 17" color monitor
- Mouse: Scroll or Optical Mouse or Touch Pad
- Keyboard: Standard 110 keys keyboard

VIII. CONCLUSION AND FUTURE SCOPE

Stroke is a critical medical condition that should be treated before it worsens. Building a

machine learning model can help in the early prediction of stroke and reduce the severe impact of the future. This paper shows the performance of various machine learning algorithms in successfully predicting stroke based on multiple physiological attributes, expert knowledge, edge detect and the combination of inaccurate information from different sources. the intend of convolution neural network is to get the appropriate classification. This project hence helps to predict the stroke risk using prediction model and provide personalized warning and the lifestyle correction message through a web application. By doing so, it urges medical users to strengthen the motivation of health management and induce changes in their health behaviors.

- FUTURE WORK

This project helps to predict thestroke risk using prediction model in older people and for people who are addicted to the risk factors as mentioned in the project. In future, the sameproject can be extended to give the stroke percentage using the output of current project. This project can also be used to findsthestroke probabilities in young people and underage people by collecting respective risk facto information'sand doctors consulting.

REFERENCES

[1] Singh, M.S., Choudhary, P., Thongam, K.: A comparative analysis for various stroke prediction techniques. In: Springer, Singapore (2020)

[2] Pradeepa, S., Manjula, K. R., Vimal, S., Khan, M. S., Chilamkurti, N., &Luhach, A. K.: DRFS: Detecting Risk Factor of Stroke Disease from Social Media Using Machine Learning Techniques. In Springer (2020).

[3] Dataset named 'Stroke Prediction Dataset' from Kaggle: https://www.kaggle.com/fedesoriano/stroke-prediction-dataset

[4] Wang, S., Li, Y., Tian, J., Peng, X., Yi, L., Du, C., Feng, C., Liang, X. (2020). A randomized controlled trial of brain and heart health manager-led mHealth secondary stroke prevention. Cardiovascular Diagnosis and Therapy, 10(5): 1192-1199. https://doi.org/10.21037/cdt-20-423

[5] Wilkinson, D.A., Daou, B.J., Nadel, J.L., Chaudhary, N., Gemmete, J.J., Thompson, B.G., Pandey, A.S. (2020). Abdominal aortic aneurysm is associated with subarachnoid hemorrhage. Journal of Neuro Interventional Surgery. https://doi.org/10.1136/neurintsurg-2020-016757

[6] Verma, A., Jaiswal, S., Sheikh, W.R. (2020). Acute thrombotic occlusion of subclavian artery presenting as a stroke mimic. Journal of the American College of Emergency Physicians Open, 1(5): 932-934. https://doi.org/10.1002/emp2.12085

[7] Yu, J., Park, S., Lee, H., Pyo, C.S., Lee, Y.S. (2020). An elderly health monitoring system using machine learning and in-depth analysis techniques on the NIH stroke scale. Mathematics, 8(7): 1-16. https://doi.org/10.3390/math8071115

[8] Xie, Y., Jiang, B., Gong, E., Li, Y., Zhu, G., Michel, P., Wintermark, M., Zaharchuk, Z. (2019). Use of gradient boosting machine learning to predict patient outcome in acute ischemic stroke on the basis of imaging, demographic, and clinical information. American Journal of Roentgenology, 212(1): 44-51. https://doi.org/10.2214/AJR.18.20260

[9] Boukobza, M., Nahmani, S., Decschamps, L., Laissy, J.P. (2019). Brain abscess complicating ischemic embolic stroke in a patient with cardiac papillary fibroelastoma - Case report and literature review. Journal of Clinical Neuroscience, 66: 277-279. https://doi.org/10.1016/j.jocn.2019.03.041

[10] S. Sung, S.M., Kang, Y.J., Cho, H.J., Kim, N.R., Lee, S.M., Choi, B.K., Cho, G. (2020). Prediction of early neurological deterioration in acute minor ischemic stroke by machine learning algorithms. Clinical Neurology and Neurosurgery, 195: 105892. https://doi.org/10.1016/j.clineuro.2020.105892