Brain Stroke Prediction Using Machine Learning and Data Science

VEMULA GEETA[1], T. MAMATHA[2], DR. ARUNA VARANASI[3], ADIMALLA PAVAN KUMAR[4], BILLA CHANDRA KIRAN[5], V. S. B. HRITHIK REDDY[6]

[1, 2] *Assistant Professor, Department of Computer Science and Engineering, Sreenidhi Institute of Science and Technology, Telangana.*

[3] *HOD, Department of Computer Science and Engineering, Sreenidhi Institute of Science and Technology, Telangana.*

[4, 5, 6] *UG Student, Department of Computer Science and Engineering, Sreenidhi Institute of Science and Technology, Telangana.*

Abstract*— One of the leading causes of mortality in contemporary society is stroke. Within a month of a stroke, at least half of the survivors will be unable to do basic tasks. Approximately 15 million individuals worldwide suffer from strokes each year, according to these data. A stroke is the second leading cause of mortality in those over 60 and the sixth highest cause of death in people aged 15 to 59, according to the World Health Organization. Over six million individuals die each year as a result of a stroke. Strokes happen to one in every six persons at some point in their lives. A person dies from a stroke every six seconds. According to the findings, high blood pressure, smoking, diabetes, and hyperlipidemia are remain the leading causes of stroke mortality. Strokes has never been able to kill as many people as the three other illnesses put together. A stroke of the brain may lead to death or long-term disability; hence it is a medical emergency. If you need immediate medical treatment, dialing 1-1-2-3 or 9-1-1 is the best option (International). Brain Attacks devastate slives around the world.

- ***13.7M newstrokeseach year.***
- ***80Mstrokesurvivorsworldwide.***
- ***5.5M deathdueto stroke each year.***
- ***1in4 people over age 25will experience stroke in their life time.***

In light of the latest statistics from the World Stroke Organization, we decided to develop a prediction model to determine whether or not an individual is suffering from a brain stroke by fine-tuning the data set obtained from the Kaggle data repository.

I. INTRODUCTION

Machine Learning, Artificial Intelligence, and Deep Learning are all revolutionizing our world, allowing us to find new ways to solve old problems and saving us time and money in the process.

As, With the rapid advancement of machine learning algorithms, which can now provide an accurate and speedy prediction result, the field of health care has gained a valuable instrument for providing individualized clinical treatment to patients suffering from different diseases. Our project's goal is to create a superior prototype. There is a good chance it will assist lower mortality rates associated with a certain disease area based on all of the available assessment measures.

II. OBJECTIVES

Detecting or Predicting whether a person is suffering from a brain stroke is the primary goal of our project, which is based on extracting patterns and insights from the data acquired from various sources. In the Exploratory Data Analysis (EDA) stage, we may also uncover useful information, such as whether men or females are more likely to suffer from this illness, what the minimum age group of those affected is, and any other permutations and combinations, in addition to the Predictive Prediction.

- Existing System:

For early screening of head injuries and brain-related diseases, the current standard is a non-contrast CT scan. Other tests that we may include are:

1. Aphysicalexam
2. BloodTests
3. Computerizedtomography(CT)scan
4. Magneticresonanceimaging(MRI)
5. carotidultrasound

6. Cerebralangiogram
7. Electrocardiogram

- Proposed System:

Based on the information provided by the patient, we have developed a system/model that may assist us recognize a patient who is suffering from brain stroke as early as feasible. Even if we expect an accurate result, we may still want to do the aforementioned tests to pinpoint the precise site of the blood cloth so that we may take the appropriate measures.

As a result, we can forecast strokes in patients, and our model might assist patients avoid stage softening, which could save both time and money.

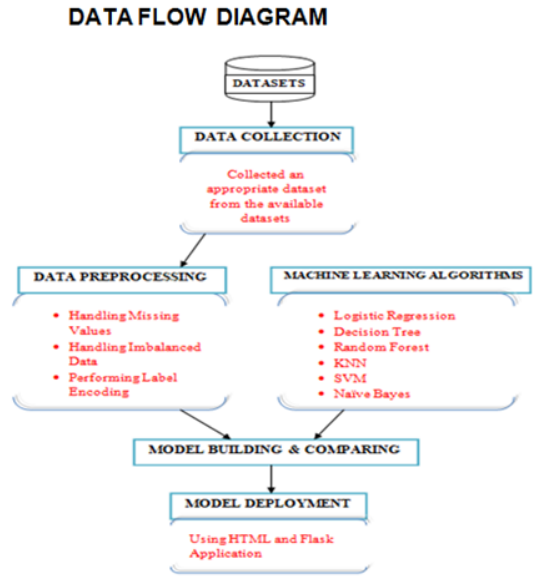
Software Requirements:

1. Google Co lab
2. Jupyter Note book
3. Libre Office
4. Python
5. MicrosoftWindows10
6. Ms Office
7. Vscode
8. Spotify

HardwareRequirements:

1. Intelcore i5 orAMDRyzen 5
2. Minimum10GBSSD/HDD
3. Ram Preferably>8for smooth execution,if not 4GB might be sufficient.
4. System Compatible with Windows/Linux/Mac Operating System.

UML Diagrams:



1. Use Case Diagram

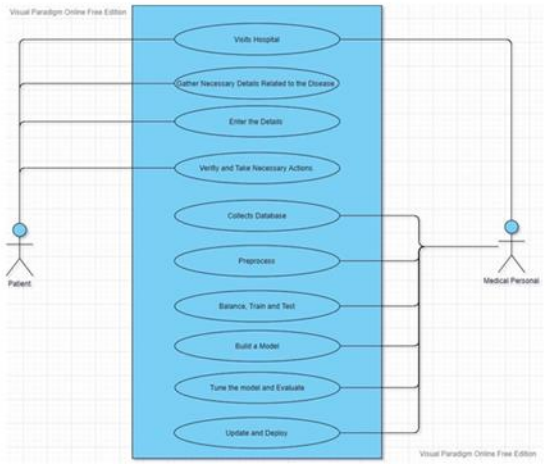


Fig 1: Use Case Diagram

2. Class Diagram

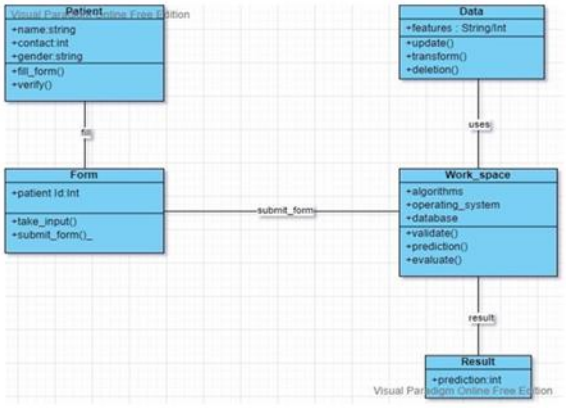


Fig 2: Class Diagram

Activity Diagram

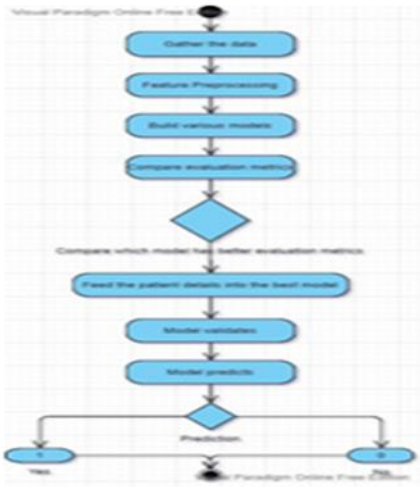


Fig 3: Activity Diagram

III. TESTING

The primary goal of testing is to demonstrate that the system is error-free and to identify any faults that may be there. Finding mistakes is the primary goal of testing. Tests let us to verify the functioning of all 41 parts, assemblies, and sunassemblies. It's done to make sure that every aspect of the program/system is operating as intended. Software requirements and user expectations may be verified by using this tool. There are a variety of tests available.

TYPES OF TESTING

Unit testing:

Unit testing has the goal of ensuring that the core logic of the program is working properly. To ensure that the program receives the right inputs and generates the necessary outputs, it is thoroughly tested. Unit testing is used to verify the source code and identify any issues. Testing of statements, branches and paths are all part of this process to ensure that the code runs correctly. A single test case may access all of the branches, as a consequence Prior to integrating all of the modules into one, it is essential to do unit testing. Unit testing is performed at the component level to ensure that a specific route of the program behaves exactly as the user needs and achieves the desired results.

Integration testing:

Integration testing is done to check that all of the components of the module have been successfully integrated and that the module performs as expected. This kind of testing is based on events. It is necessary to do integration testing in order to verify that the different components, which had previously worked flawlessly together, will continue to do so. Integration testing is used to identify and correct errors that may have occurred during the integration process, as well as to verify that all of the components work together harmoniously. The user's chosen module should be created by combining the various components.

System Testing:

In order to be certain that the integrated system meets the needs of the end user, it is necessary to conduct system testing. To do this, you must first verify the set-up is correct. This is an integration test for a system that is configuration-driven. Rather, it focuses on how the many processes interact with one another. Integration points and pre-driven process connections are also examined.

AcceptanceTesting:

As a vital element of every project, acceptance testing necessitates the cooperation of the end user. If all functional, nonfunctional and user requirements are fulfilled by the system then it is a successful system implementation.

S.NO	INPUT	OUTPUT	RESULT
Test Case 1 (Decision Tree testing of Dataset)	The Patient gives the Data to predict the stroke.	An output is predicts stroke or not.	A result predicts stroke or not..
Test Case 2 (Decision tree tuning for Accuracy)	The patient enters data in the Database to predict the stroke.	An output predicts stroke or not.	A result using Decision tree predicts stroke or not

Test Case 3 (Logical Regression testing)	The patient enters data in the Database to predict the stroke.	An output predicts stroke or not.	A result using Logical Resgression predicts stroke or not
Test Case 4 (Logical Regression tuning)	The patient enters data in the Database to predict the stroke.	An output predicts stroke or not.	A result using Logical Resgression predicts stroke or not
Test Case 5 (Logical Regression testing)	The patient enters data in the Database to predict the stroke.	An output predicts stroke or not.	A result using Logical Resgression predicts stroke or not
Test Case 6 (Random Forest testing)	The patient enters data in the Database to predict the stroke.	An output predicts stroke or not.	A result using Random Forest predicts Brain stroke or not
Test Case 7 (Random forest testing)	The patient enters data in the Database to predict the stroke.	An output predicts stroke or not.	A result using Random Forest predicts Brain stroke or not
Test Case 8 (Random Forest Tuning)	The patient enters data in the Database to predict the stroke.	An output predicts Brain stroke or not.	A result using Random Forest predicts Brain stroke or not
Test Case 9	The patient enters data in the Database to predict the stroke.	An output predicts Brain stroke or not.	Predicts the chances of Brain Stroke
Test Case 10	The patient enters data in the Database to predict the stroke.	An output predicts Brain stroke or not.	Predicts the chances of Brain Stroke
Test Case 11	The patient enters data in the Database to predict the stroke.	An output predicts Brain stroke or not.	Predicts the chances of Brain Stroke
Test Case 12	The patient enters Health data in the Database to predict the stroke.	An output predicts Brain stroke or not.	Predicts the chances of Brain Stroke
Test case 13	The patient enters Health data in the Database to predict the stroke.	An output predicts Brain stroke or not.	Predicts the chances of Brain Stroke
Test case 14	The patient enters Health data in the Database to predict the stroke.	An output predicts Brain stroke or not.	Predicts the chances of Brain Stroke
Test Case 15	The patient enters Health data in the Database to predict the stroke.	An output predicts Brain stroke or not.	Predicts the chances of Brain Stroke
Test Case 16	The patient enters Health data in the Database to predict the stroke.	An output predicts Brain stroke or not.	Predicts the chances of Brain Stroke
Test Case 17	The patient enters Health data in the Database to predict the stroke.	An output predicts Brain stroke or not.	Predicts the chances of Brain Stroke
Test Case 18	The patient enters Health data in the Database to predict the stroke.	An output predicts Brain stroke or not.	Predicts the chances of Brain Stroke
Test Case 19	The patient enters Health data in the Database to predict the stroke.	An output predicts Brain stroke or not.	Predicts the chances of Brain Stroke
Test Case 20	The patient enters Health data in the Database to predict the stroke.	An output predicts Brain stroke or not.	Predicts the chances of Brain Stroke

OUTPUT SCREENS:

Fig 4: Importing Libraries

Importing and Skimming of Dataset.

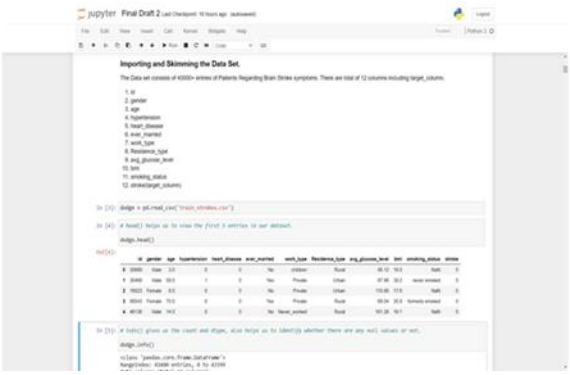


Fig 5: Important and Skimming of dataset

Exploring Target Variable.

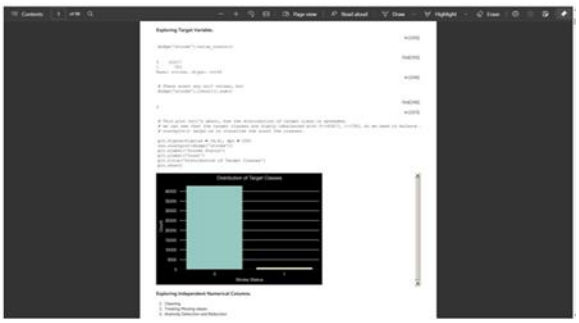


Fig 6: Exploring Target Variable

Exploratory Data Analysis.

Fig 7 : Exploratory Data Analysis

Sample Auto Viz Charts.

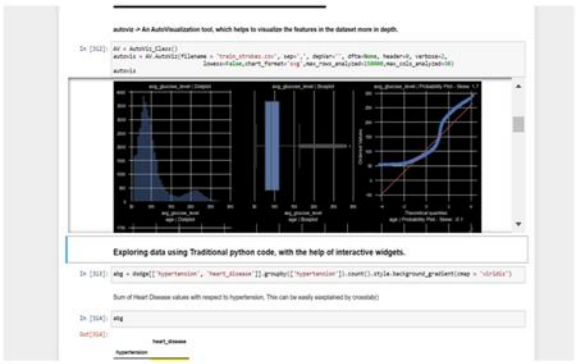


Fig 8 : Sample Auto Viz Charts

IV. CONSLUSION

Ultimately, we want to learn more about how artificial intelligence and machine learning are rapidly altering the world around us.

With the high dose of radiation that CT scans provide to patients, we set out to create a model that may potentially restrict its use in early screening for head traumas and brain diseases. future.

Those with a high False Positive Rate, in which the individual is not suffering from a stroke but is nonetheless exposed to a costly and time-consuming diagnostic process, may benefit from this technique.

V. FUTURE SCOPE

Due to the fact that we have just constructed a prototype model that accepts input and outputs a result in the formatof1/0, there are still many aspects of this project that need more examination, which we want to undertake in the future.

REFERENCES

[1] *https://www.world-stroke.org/about-wso/annual-reports*

[2] *https://www.medicinenet.comstroke_symptoms_and_treatment/article.htm*

[3] *https://towardsdatascience.com/data-science/home*

[4] *https://medium.com/*

[5] *https://www.analyticsvidhya.com/*

[6] *https://www.youtube.com/results?search_query=krish+naik+data+science*

[7] *https://scholar.google.com/*

[8] *https://www.kaggle.com/*

[9] *https://trainings.internshala.com/*

[10] *https://www.theedgemarkets.com/article/budget-2021-healthcare-measures-welcomed-fall-short*