

**A
Project Report
On
“Stroke Prediction”**

Prepared by
Harikesh Prajapati (18DCS093)
Darshan Raval (18DCS099)
Rajvi Shah (18DCS109)

Under the guidance of
Mr. Minal Patel,
Assistant professor

A Report Submitted to
Charotar University of Science and Technology
for Partial Fulfillment of the Requirements for the
7th Semester Software Group Project-V (CS448)

Submitted at



**CSE
DEPSTAR
At: Changa, Dist: Anand – 388421
November 2021**



Certificate

This is to certify that the report entitled “**Stroke Prediction**” is a bonafied work carried out by **Harikesh Prajapati (18DCS093)**, **Darshan Raval (18DCS099)** and **Rajvi Shah (18DCS109)** under the guidance and supervision of **Assistant Prof. Mr. Minal Patel** for the subject **CS448-Software Group Project-V (CSE)** of 7th Semester of Bachelor of Technology in **DEPSTAR** at Faculty of Technology & Engineering – CHARUSAT, Gujarat.

To the best of my knowledge and belief, this work embodies the work of candidate himself, has duly been completed, and fulfills the requirement of the ordinance relating to the B.Tech Degree of the University and is up to the standard in respect of content, presentation and language for being referred to the examiner.

Mr. Minal Patel
Assistant Professor
CE
DEPSTAR, Changa, Gujarat.

Dr. Amit Ganatra
Principal, DEPSTAR
Dean, FTE
CHARUSAT, Changa, Gujarat.

**Devang Patel Institute of Advance Technology And Research At: Changa, Ta.
Petlad, Dist. Anand, PIN: 388 421. Gujarat**

ACKNOWLEDGEMENT

We have found this rare opportunity to evince a word of thanks to all those who played a key role in the successful completion of our project. We sincerely thank our Head of Department **Dr. Amit Ganatra** Sir for giving the chance as well as support for all the time being. And his able guidance and continuous encouragement made us work in all the challenges during project development. We express deep gratitude to **Assistant Prof. Mr. Minal Patel**, internal project guide from Faculty of Engineering, DEPSTAR for their valuable suggestions, help and moral support.

Harikesh Prajapati (18DCS093)

Darshan Raval (18DCS099)

Rajvi Shah (18DCS109)

ABSTRACT

A large number of people lose their life due to stroke and it is increasing in developing countries. There are several stroke risk factors that regulate different types of stroke. Predictive algorithms help to understand the relation between these risk factors to types of strokes. We have used several machine learning algorithms to detect the type of stroke that can occur in a patient or already occurred from their clinical report and statistical data. We have used a stroke dataset by collecting data from various sources. Then the dataset was processed to be used with the machine learning algorithms.

TABLE OF CONTENTS

Acknowledgement.....	3
Abstract.....	4
Chapter 1 Project definition.....	7
Chapter 2 Description.....	8
2.1) Libraries Used.....	8
2.2) Approach.....	8
Chapter 3 Software and Hardware Requirement.....	9
Chapter 4 Algorithms Used.....	10
Chapter 5 Screenshot of Project.....	11
Chapter 8 Outcomes.....	12
Chapter 10 Conclusion.....	13
Chapter 11 References.....	14

LIST OF FIGURES

Fig. 4.1 Algorithm vs Accuracy Graph.....	10
Fig. 5.1 Entering input values.....	11
Fig. 5.2 Detection of Stroke.....	11

CHAPTER-1 PROJECT DEFINITION

Stroke Detection is an application that can be used to improve patients' health through early detection and treatment. We will be using different Machine Learning algorithms for prediction. Once the model is ready and gives accurate results, we will deploy the model into a web application. The application will take various parameters as input from the user such as gender, age, hypertension, heart disease, etc. and based on that prediction will be done.

CHAPTER-2 DESCRIPTION

2.1 Libraries Used:

➤ The following libraries were used in creating this project:

- 1) **Pandas** - For data pre-processing
- 2) **Matplotlib** - For data visualization
- 3) **Scikit Learn** - For building Prediction Models
- 4) **Flask** - For deployment of model into a web application

2.2 Approach:

➤ The following approach was used:

- 1) **Data pre-processing** - Importing libraries, exploratory data-analysis.
- 2) **Label Encoding and Normalisation** - Converting the labels such as gender into numeric form, changing the values of numeric columns to a common scale.
- 3) **Building the Model** - Dividing the dataset into training and testing data and apply the algorithms.
- 4) **Deployment** - Working on Front-end and web deployment using Flask.

CHAPTER-3 SOFTWARE AND HARDWARE REQUIREMENTS

➤ **Hardware:**

- 1) Minimum 8GB RAM for better performance.
- 2) i5 8th Gen or above processor.

➤ **Software:**

- 1) Anaconda/Pycharm as python interpreter.

CHAPTER-4 ALGORITHMS USED

We have implemented the following algorithms.

1. Decision Tree
2. Logistic Regression
3. kNN
4. Random Forest
5. SVM

The graph below shows accuracy of all the 5 algorithms.

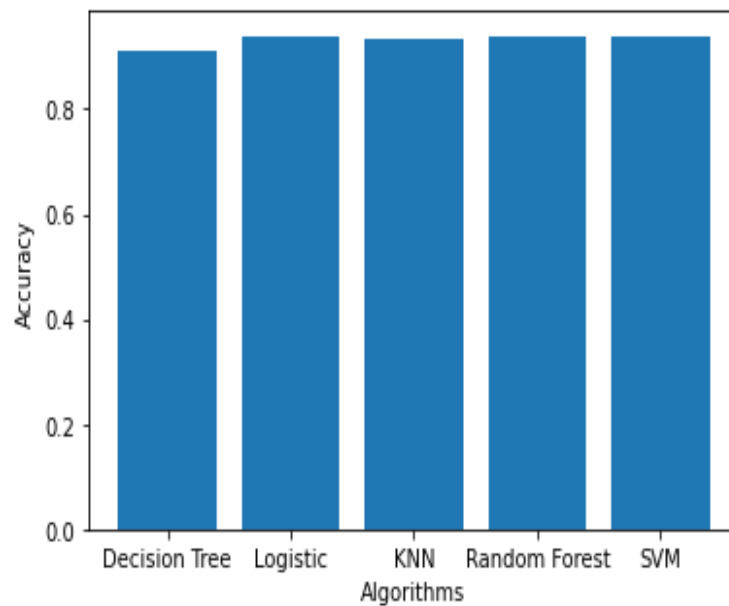


Fig. 4.1 Algorithm vs Accuracy Graph

CHAPTER-5 SCREENSHOTS OF PROJECT

Please enter your details in the following form if you want to know whether you are prone to suffer from a stroke.

Gender	1
Age	67
Hypertension	1
Heart Disease	1
Ever Married	1
Work Type	2
Residence Type	1
Average Glucose Level	228.67
BMI	36.6
Smoking Status	1

Submit

Fig. 5.1 Entering input values



You have been diagnosed with Stroke Risk

The algorithm has diagnosed you with Stroke Risk based on your inputs. Please consult a Doctor.

Fig. 5.2 Detection of Stroke

CHAPTER-6 OUTCOMES

- This project was a very good way to probe into the field of Machine Learning. Supervised Learning Models were effectively learnt and implemented.
- Also, we have learnt many things from this project including the co-ordination for working in a team.
- The importance of time-bound and coordination was realized.

CHAPTER-7 CONCLUSION

This project classifies a sufficiently large dataset of stroke-attacked patients accurately. Amongst Decision Tree, Logistic Regression, kNN Random Forest and SVM, the algorithm with a better accuracy will be used. The model can help people with a cautionary indication of being affected by stroke. Healthcare industries generate huge amounts of complex data about patients, hospitals resources, disease diagnosis, electronic patient records, medical devices, etc. Thus, this application provides an easier approach.

CHAPTER-8 REFERENCES

- https://www.researchgate.net/publication/338361726_Detection_of_Stroke_Diagnosis_using_Machine_Learning_Algorithms
- <https://ieeexplore.ieee.org/document/8944689>
- <https://www.youtube.com/watch?v=UtgrhBr3kTw>
- <https://www.youtube.com/watch?v=hJ2sKPj5Xn4>
- <https://www.youtube.com/watch?v=Qr4QMBUPxWo>