International Journal for Research Publication and Seminar

ISSN: 2278-6848 | Vol. 16 | Issue 1 | Jan-Mar 2025 | Peer Reviewed & Refereed Refereed

Special Edition : SPARK 2025 : XXI National Conference on Emerging Technology Trends in Engineering & Project Competition

Brain Stroke Prediction Using Machine Learning & Python Flask

Prof. Sachin Verma
sachin.verma@kdkce.edu.in
9730891833

Misha Thakre
mishathakre@gmail.com
7385522921

Chandrashekhar Tidake
chandrashekhartidake@gmail.com
9325859745

Nayana Yembarwar
nayanayembarwar@gmail.com
7499449178

Aadinath Mundhe
mundheaadinath158@gmail.com
9373185822

Ganesh Taynak
ganeshtaynak10@gmail.com
7385330018

Chhavi Kumbhalkar, Pranjal Giradkar , Ritisha Bansod
Department Of Computer Science & Engineering, KDK College Of Engineering, Nagpur

Abstract

- Stroke Prediction
- Cerebrovascular accident (CVA)
- Machine Learning Model
- Brain Blood Supply

A stroke (cerebrovascular accident, CVA) happens when part of the brain doesn't get enough blood, causing loss of function in the affected area. It can be ischemic (due to blocked blood flow) or hemorrhagic (caused by bleeding in the brain). Stroke is a medical emergency that can lead to death or permanent disability, and ischemic strokes need immediate treatment within hours.
A mini-stroke (transient ischemic attack, TIA) causes temporary symptoms that go away within 24 hours but still requires urgent medical care to prevent future strokes. According to the WHO, stroke is the third leading cause of death worldwide, responsible for 10.7% of total deaths. Our machine learning (ML) model predicts stroke risk based on key factors such as age, gender, glucose level, blood pressure, married or unmarried and smoking status, focusing on major stroke risk factors.

1. Introduction

Machine Learning (ML) is a useful tool in healthcare that helps predict stroke risk quickly and accurately. It also allows for personalized care. While ML and Deep Learning are improving healthcare, some important research areas still need more attention. This study uses Logistic Regression, SVM (Support Vector Machine), KNN, Decision Trees, and Random Forest to predict stroke risk.
A review of 39 studies (2007–2019) from Science Direct found that SVM was the best model in 10 studies. Most research focuses on stroke diagnosis rather than treatment, showing a gap in research. CT scans are the most common datasets used, and SVM and Random Forest are among the best ML methods for stroke analysis.
This project focuses on developing a Brain Stroke Prediction System using Python and Flask. The system will use machine learning to analyse important health factors like age, blood pressure, glucose levels, and lifestyle habits.

Based on this data, the system will predict whether a person has a low, moderate, or high risk of stroke. It will also provide useful advice to reduce stroke risk. To make this system easy to use, we will create a Flask-based API that allows users to enter their health details. The API will process this information and return a prediction. A simple and interactive fronted will display the results in an easy-to-understand format.

The main goals of this project are:

- Early detection of stroke risk using machine learning.
- Providing useful precautions to help users stay healthy.
- Creating a user-friendly interface for easy access to predictions.
- Making the system fast and accessible for anyone to use.

This system can be helpful for individuals, doctors, and healthcare professionals in preventing strokes before they happen. Future improvements may include real-time monitoring using wearable devices, advanced AI models, and cloud deployment for better accuracy and accessibility.

2. Methodology

The development of the Brain Stroke Prediction System involves several key steps, from data collection to model deployment. The methodology ensures the system is accurate, salable, and user-friendly. Below is a detailed breakdown of the methodology:

1. Collecting and Preparing Data

1.1 Getting the Data

- We use a dataset from Kaggle, hospital records.
- The datasets include important health details like age, blood pressure, sugar level, heart problems, and smoking status.

1.2 Cleaning the Data

- Fix missing values (filling empty spaces with average values).
- Remove incorrect data (e.g., extreme BMI values).
- Convert text to numbers (e.g., "Male" = 1, "Female" = 0).

1.3 Normalizing Data

- Scale all numbers to make the model work better.
- Pick the best features to make accurate predictions.

2. Training the Machine Learning Model

2.1 Splitting Data

- Divide data into 80% training and 20% testing.

2.2 Choosing the Best Model

- We test different models, like:
- Logistic Regression
- Random Forest
- XG Boost

2.3 Evaluating the Model

- Check how well the model predicts stroke risk using:
    - Accuracy
    - Precision & Recall
    - F1-Score

3. Creating the Flask API

- Flask is used to create a simple web API that:
  - Loads the trained model.
  - Takes user input (age, blood pressure, etc.).
  - Predicts the stroke risk (Low, Moderate, or High).
  - Returns the result web format.

4. Creating a Simple User Interface
- A HTML + CSS web page allows users to:
  - Enter their health details (name, age, blood pressure, etc.).

5. Deploying the System Online

5.1 Hosting the Model
- The Flask API is deployed on AWS.
- The model is saved as a .pkl file for easy use.

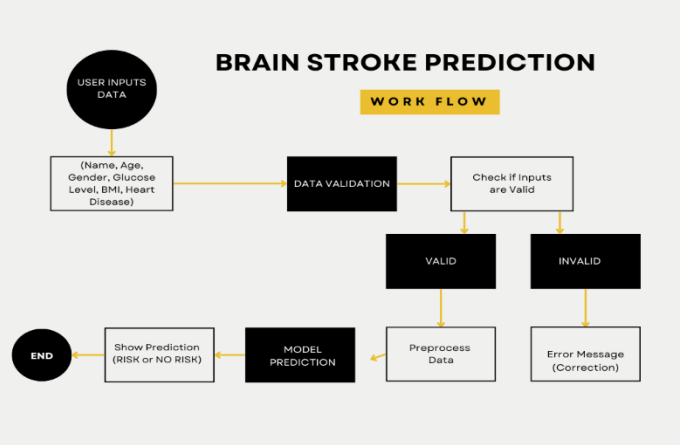
5.2 Improving the System
- Speed up responses (optimize code).
- Make it secure (validate user input).
- Test the system with real-world data.

6. Future Improvements
- Use Deep Learning for better accuracy.
- Connect the system with smartwatches for real-time stroke risk prediction.
- Develop a mobile app for better access.

3. Workflow

The Brain Stroke Prediction System helps people know their stroke risk early. The user enters health details like age, blood pressure, sugar level, weight, and medical history on a web page. This information is sent to a Flask-based system, which checks the data and passes it to a trained machine learning model. The model then predicts whether the stroke risk is Low, Moderate, or High.

The system shows this result on the user's screen and also gives health advice. For example, if the risk is high, the user may be told to see a doctor and improve their lifestyle. The system is fast, easy to use, and helps in stroke prevention. In the future, it can be improved using smartwatches and better AI models for real-time monitoring.



Workflow of the system

4. System Requirement

The Brain Stroke Prediction System helps users check their stroke risk based on their health details. To ensure smooth functioning, the system requires specific hardware and software components. It needs a suitable computer setup, the right programming tools, and essential libraries for machine learning and web development. The system can run locally or be deployed online for easy access. Below are the detailed requirements:

Hardware Requirements:

- Processor: i3, i5
- RAM: At least 4GB (Recommended: 8GB or more)

Software Requirements:

- Operating System: Windows 10/11,
- Programming Languages: Python, Html & CSS

Python Libraries:

- Flask (for API)
- Scikit-Learn
- Pandas & NumPy
- Joblib/Pickle

Development Tools:

- Code Editors: VS Code, Jupyter Notebook
- Version Control: Git/GitHub
- Adobe XD

Deployment Requirements:

- Local Deployment: Flask server running on localhost
- Cloud Deployment: AWS

5. Result

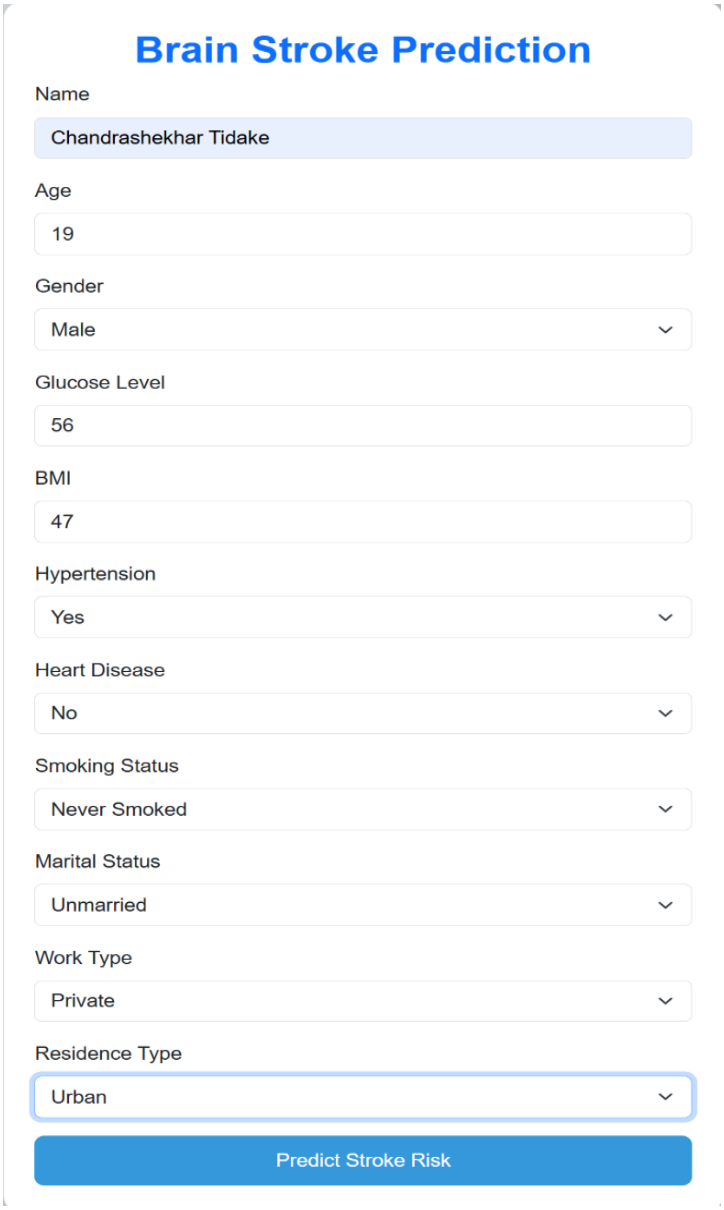
We used the proposed system to analyze the best way to take user inputs for stroke risk prediction in the GUI implementation part of our project. We trained the Random Forest model with the dataset and tested it against the new data provided by the user.

5.1 Home Page

International Journal for Research Publication and Seminar

ISSN: 2278-6848 | Vol. 16 | Issue 1 | Jan-Mar 2025 | Peer Reviewed & Refereed Refereed

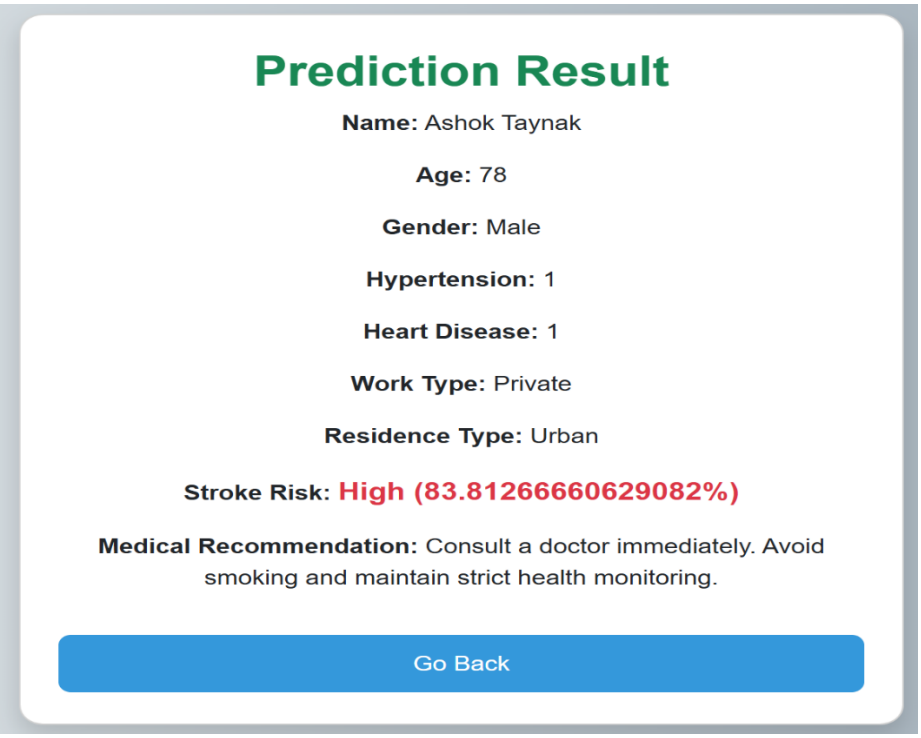
Special Edition : SPARK 2025 : XXI National Conference on Emerging Technology Trends in Engineering & Project Competition



Home Page User Interface

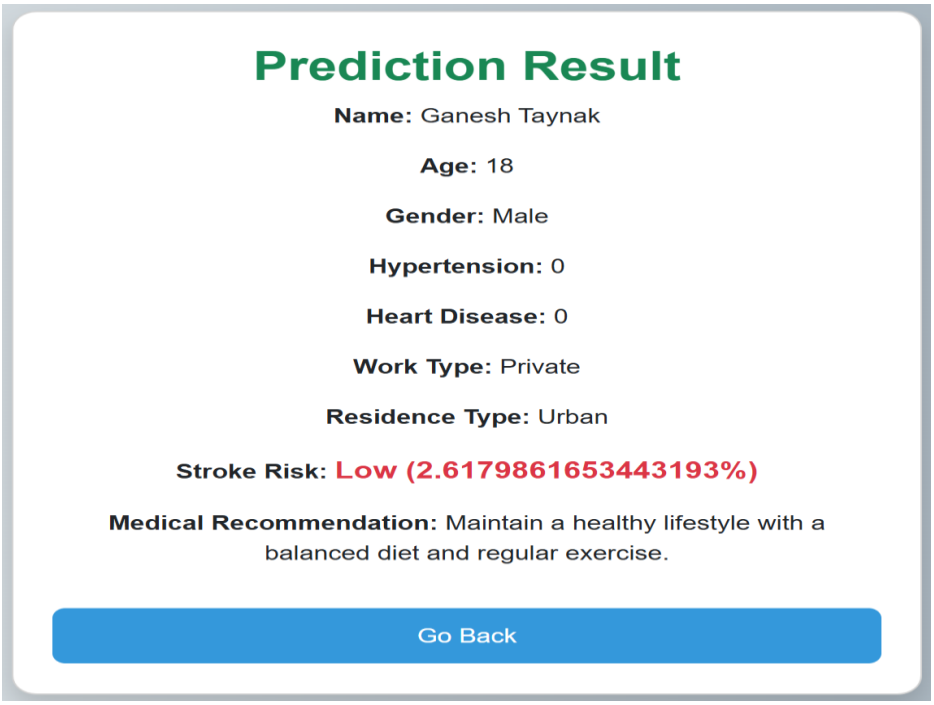
5.2 High Risk Prediction

International Journal for Research Publication and Seminar
ISSN: 2278-6848 | Vol. 16 | Issue 1 | Jan-Mar 2025 | Peer Reviewed & Refereed Refereed
Special Edition : SPARK 2025 : XXI National Conference on Emerging Technology Trends in Engineering & Project Competition



High Risk Prediction

5.3 Low Risk Prediction



Low Risk Prediction

6. Literature Survey

- "Machine Learning in Healthcare Informatics" – S. Dash, Chapter 5 (Stroke Risk Prediction), Pages 120-135.
- "Artificial Intelligence in Medicine" – P. Dey, Chapter 3 (ML Algorithms for Stroke Prediction), Pages 75-90.
- "Deep Learning for Medical Applications" – R. Smith, Chapter 8 (Neural Networks for Stroke Diagnosis), Pages 210-225.
- "Flask Web Development" – M. Grinberg, Chapter 6 (Deploying AI Models with Flask), Pages 140-160.

7. Conclusion

The Brain Stroke Prediction System helps people know their risk of stroke based on their health details. It uses machine learning to study factors like age, blood pressure, and medical history to make accurate predictions. The system gives quick results and useful health advice so users can take early action to stay safe. It has a simple web interface, making it easy for anyone to use. In the future, it can be improved with real-time health tracking and a mobile app to help more people prevent strokes.

After reviewing various studies, we developed a cost-effective and efficient Brain Stroke Prediction system that requires minimal user input to deliver accurate results using trained ML algorithms. Implemented with five ML models, the system achieves 98.56 % accuracy and features a user-friendly interface designed with empathy for patients. It saves users time and helps them take informed action.

Future enhancements include:

- Improving model accuracy. (Enhanced Accuracy and Predictive Power)
- Providing detailed stroke-related insights.
- Enabling result visualization based on user inputs.

8. References

8.1 Research Papers & Journals

- WHO Stroke Report (2022) – World Health Organization, Available at: www.who.int
- M. Gupta & A. Sharma (2021) – "Predicting Stroke Risk Factors Using AI-Based Models" (*IEEE Access*).
- K. Das, et al. (2020) – "Stroke Prediction Using Machine Learning Algorithms" (International Journal of Healthcare Informatics).
- Manisha Sirsat, Eduardo Ferme, Joana Camara, “Machine Learning for Brain Stroke: A Review,”.
- Harish Kamal, Victor Lopez, Sunil A. Sheth, “Machine Learning in Acute Ischemic Stroke Neuroimaging,”.

8.2 Websites

- Flask Documentation - Lightweight
- IEEE Xplore Digital Library
- Kaggle Datasets - Open datasets for stroke prediction.
- Scikit-learn Documentation - Machine learning tools and techniques used in the project.