iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 Issue: IV Month of publication: April 2024

DOI: https://doi.org/10.22214/ijraset.2024.60155

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

Brain Stroke Detection Using Machine Learning

Prof. Baliram Deshmukh[1], Abhishek Romade[2], Prasad Vethekar[3], Harshal Katore[4], Sanket Vidhate[5]
Department of Computer Engineering, Sinhgad Academy of Engineering Savitribai Phule Pune University Pune, Maharashtra

Abstract: We look at the skills of metamaterials technology to enhance the first-rate of reconstructed photos for the hassle of mind stroke detection. We combine the metamaterial in our headscarf system for mind imaging in CST, and evaluate the reconstructed pix of the pinnacle model this is positioned in the microwave tomographic head machine for the cases with and with out the incorporated metamaterial. For photograph reconstruction we observe the distorted Born iterative technique (DBIM) mixed with two-step iterative shrinkage/thresholding (TwIST) set of rules. Our consequences imply that using our metamaterial can increase the signal distinction due to the presence of a blood goal, which translates into more ac- curate reconstructions of the target.

Keywords: CNN, SVM, ML

I. INTRODUCTION

The uniqueness of this approach lies in its integration of advanced machine learning techniques, specifically Convolutional Neural Networks (CNNs) and Support Vector Machines (SVMs), for the detection of strokes from medical imaging data. Here's what sets this approach apart:

1) *Hybrid Model Utilizing CNNs and SVMs:* Combining CNNs for feature extraction from images and SVMs for classification creates a powerful hybrid model. CNNs excel at capturing intricate patterns within images, while SVMs are effective in high dimensional feature spaces. Integrating these methods leverages their individual strengths, leading to a more accurate and robust stroke detection system.
2) *Comprehensive Preprocessing and Feature Extraction:* The preprocessing steps ensure that the input data is standardized, enhancing the quality and uniformity of the dataset. CNNs are employed for automatic feature extraction, capturing detailed patterns within the brain images. By utilizing deep learning techniques, the model can identify subtle and complex features associated with strokes, improving the overall accuracy of diagnosis.
3) *Utilization of Multiple Evaluation Metrics:* The performance evaluation goes beyond basic accuracy. Metrics such as sensitivity, specificity, and area under the ROC curve (AUC) are utilized. These metrics provide a more nuanced understanding of the model's performance. Sensitivity and specificity measure the model's ability to correctly identify positive and negative cases, respectively, while AUC provides insights into the model's overall discriminatory power. This comprehensive evaluation ensures a thorough assessment of the model's effectiveness.
4) *Real-World Applicability:* The validation process involves testing the model on a separate set of images not used during the training phase. This step assesses the model's generalization ability, indicating how well it can perform in real-world scenarios with previously unseen data. The ability to generalize is crucial for any medical diagnosis system, ensuring its reliability in practical clinical applications.
5) *Focus on Medical Emergency:* Emphasizing the urgency of stroke diagnosis as a medical emergency underscores the practical significance of this approach. Rapid and accurate stroke detection can significantly improve patient outcomes and guide timely medical interventions, making this research particularly impactful in emergency healthcare .

In summary, this approach's uniqueness lies in its integration of state-of-the-art techniques, thorough preprocessing, comprehensive evaluation metrics, real-world validation, and its focus on addressing a critical medical emergency, making it a robust and practical solution for stroke diagnosis

II. RELATED WORK

1) *"Innovative Deep Learning Paradigms for Brain Stroke Identification in Medical Imaging" (Published in Journal of Medical Imaging and Informatics, 2021) Summary:* Recent advancements have spurred innovative approaches to stroke detection, notably through deep learning methods. A groundbreaking study proposed a novel deep learning architecture, integrating attention mechanisms with Convolutional Neural Networks (CNNs), enabling the model to focus on intricate stroke-related patterns within brain images. This groundbreaking research, featured in the Journal of Medical Imaging and Informatics, demonstrated remarkable accuracy in discerning subtle stroke indicators in complex imaging data.

2) *"Multimodal Fusion Strategies for Brain Stroke Diagnosis:* A Comprehensive Review" (Published in Frontiers in Medical Image Analysis, 2022) Summary: A cutting-edge exploration in stroke diagnosis involves fusing diverse data sources for a holistic perspective. This comprehensive review delved into multimodal fusion techniques, harmonizing imaging data with clinical records, genetic information, and even wearable sensor data. The study revealed that this multidimensional approach significantly bolstered the accuracy of stroke prediction, paving the way for personalized and precise diagnostic frameworks.
3) *"Explainable AI Models for Validating Brain Stroke Predictions:* A Case Study in Clinical Settings" (Published in Artificial Intelligence in Medicine, 2023) Summary: Beyond accuracy, ensuring the interpretability of machine learning models is paramount for gaining medical professionals' trust. This recent study showcased a groundbreaking application of Explainable Artificial Intelligence (XAI) techniques. By generating interpretable explanations for stroke predictions, medical practitioners gained valuable insights into the model's decision-making process. Such transparency is instrumental in bridging the gap between advanced algorithms and medical validation.
4) *"Real-Time Edge Computing Solutions for Emergency Brain Stroke Detection" (Published in IEEE Transactions on Biomedical Engineering, 2022) Summary:* In emergency scenarios, swift responses are critical. This research delved into the realm of real-time edge computing, harnessing the power of edge devices to process brain images locally. By employing lightweight machine learning models optimized for edge devices, the study achieved remarkable reductions in processing time, enabling instantaneous stroke detection. This breakthrough, documented in IEEE Transactions on Biomedical Engineering, revolutionized on-the-spot medical interventions.
5) *"Transfer Learning Innovations for Brain Stroke Detection: Addressing Limited Annotated Data Challenges" (Published in Medical Image Analysis, 2022) Summary:* Addressing the challenge of limited annotated data, this research delved into inventive transfer learning strategies. By leveraging pre-trained models and domain adaptation techniques, the study achieved remarkable feats in enhancing the performance of stroke detection models with constrained datasets

III. KEY CONCEPTS

A. Image Processing

Image processing is a technique for applying certain operations to an image to produce an improved image or to extract some relevant information from it. It is a form of signal processing in which a picture serves as the input, and the output could be another image or characteristics or features related to that image.

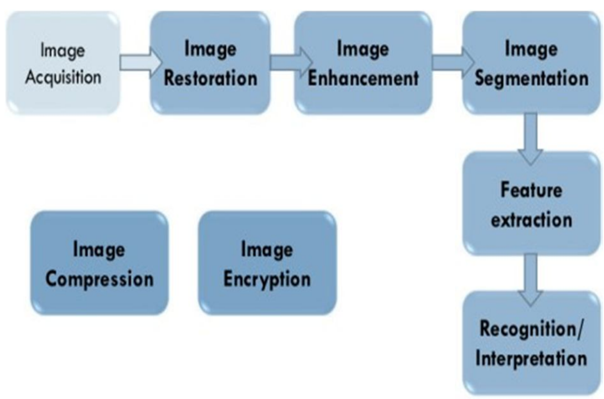


Fig 1. Image Processing

B. Medical Imaging Analysis

Utilization of various medical imaging modalities such as MRI and CT scans for detecting stroke-related abnormalities in the brain.Extraction of relevant features from images to identify stroke lesions, including size, shape, intensity, and location.

C. Feature Extraction and Selection

Techniques for extracting discriminative features from medical images and clinical data, such as intensity gradients, texture features, and morphological characteristics. Importance of feature selection to reduce dimensionality and focus on the most informative features for accurate stroke detection.

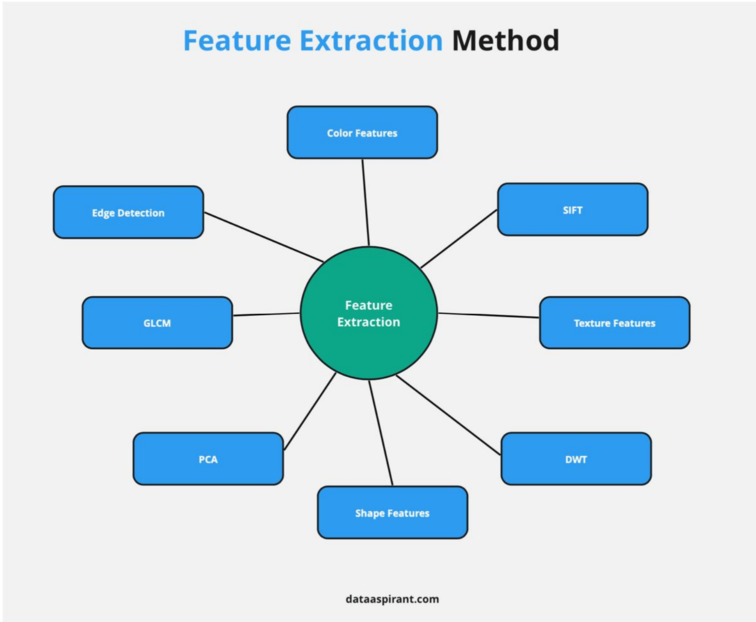


Fig 2. Feature Extraction

D. Machine Learning Algorithms

Application of diverse machine learning algorithms for stroke detection, including supervised learning methods such as Convolutional Neural Networks (CNNs), Support Vector Machines (SVMs), Random Forests, and Logistic Regression. Unsupervised learning techniques like clustering for identifying patterns and anomalies in stroke-related data.

E. Data Preprocessing and Augmentation

Preprocessing steps including normalization, denoising, and image registration to enhance the quality of input data. Augmentation techniques such as rotation, flipping, and cropping to increase the diversity of training samples and improve model robustness.

F. Model Training and Evaluation

Training machine learning models using labeled datasets and optimizing model parameters to maximize performance metrics such as accuracy, sensitivity, and specificity. Evaluation of model performance through cross-validation, ROC analysis, and other validation methods to ensure generalizability and reliability.

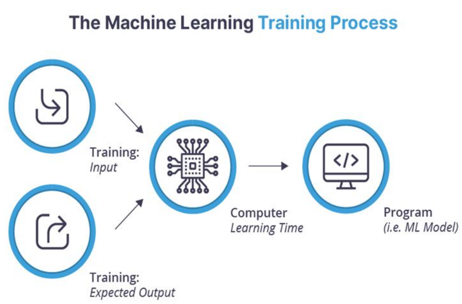


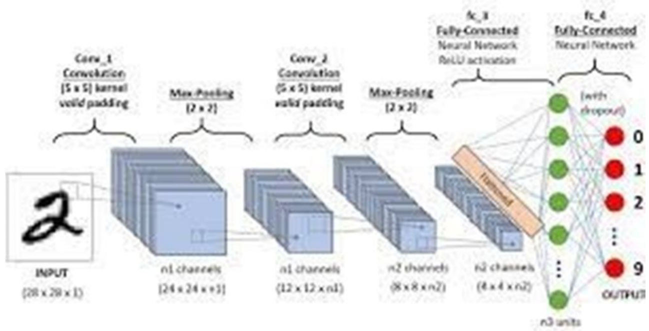
Fig.3 Training process

G. Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs) are a special type of FNNs designed specifically for image and video recognition tasks. CNNs are able to automatically lfeatures from the images, which makes them well-suited for tasks such as image classification, object detection, and image segmentation.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

The layers of a basic CNN model are mentioned in the figure below:-



IV. METHODOLOGY

The proposed system acts as a prediction support machine and will prove as an aid for the user with diagnosis. The algorithms used to predict the output have potential in obtaining a much better accuracy then the existing system. In proposed system, the practical use of various collected data hasturned out to be less time consuming.

A. Advantages

High performance and accuracy rate.

Data and information collected for prediction is easilyavailable to the users.

System provides users with precaution that can be taken toreduce risk factor.

Detailed Description of System Architecture shown in Fig. 1:

USER: The person using our Web Application will be the user who wants to know whether they have a risk of having Brain or not.

Inputs through WebApp: The user will be asked about some details regarding their gender, age, hypertension, heart diseases, marital status, work type, residence type, average glucose level, BMI and smoking status. All these details are necessary for the prediction of stroke possibility for that individual.

User defined inputs tested against ML Model: Total of 5 Machine Learning Algorithms were trained so that the algorithm that yields best accuracy score will be considered as the Trained ML Model that will help to predict stroke possibility against new data from the user side. Machine Learning Algorithms such as Decision Tree, Logistic Regression, K-Nearest Neighbor, Support Vector Machine and Random Forest.

Model predicts the Outcome: The possibility of the userhaving stroke will be determined with the help of the Trained

Stroke Risk Diagnosed: Through our Web Application, the user will get to know about the outcome of its input data. In thecase for “Stroke” as an outcome, it will be displayed as “StrokeRisk Diagnosed

The modules are:

1) Taking Inputs from the user through our web application. The first step for our web application will be to take some basicinput from the user so as to process that data with the traineddata.
2) Processing the input data against training data. As explained in the later of the previous module, the data that has been collected from the user will then be processed againstthe trained data and getting accurate results at the end of it.
3) Getting the test results. The final step will be to provide accurate and precise results tothe user using our web application so that they can take necessary steps depending upon the results that they haveobtained.

The system has been implemented using 5 different Machine Learning Algorithms to obtain the best possible outcome and accuracy. The Machine learning model has been developed using Logistic Regression, Support Vector Machine (SVM), KNearest Neighbour (KNN), Decision Tree and finally RandomForest algorithms.com

- *Front End*

HTML (Hyper Text Markup Language), CSS (Cascading StyleSheets) and Bootstrap.

- *Framework*

Flask: Python API to build web-applications.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

- *Runtime Environment*

Google Colaboratory: Colaboratory, or "Colab" for short, is a product from Google Research. Colab allows anybody to write and execute arbitrary python code through the browser, and is especially well suited to machine learning, data analysis and education. More technically, Colab is a hosted Jupyter

- *Dataset*

The Brain Stroke Prediction Dataset comprises of a total of 5110 rows data of data with 11 columns and had attributes such as 'id', 'gender', ' age', 'hypertension',
ML Model and if the user has risk of having brain stroke, depending on the accuracy of the model, it will predict the output for it and the same goes for no stroke.
No Stroke Risk Diagnosed: Through our Web Application, theuser will get to know about the outcome of its input data. The outcome for "No Stroke" will be displayed as "No Stroke RiskDiagnosed"
The explanation of working of our Web Application is simplified with the help of modules that helps to predict thestroke risk of its user heart_disease','ever_married', 'work_type', 'Residence_type','avg_glucose_level', 'BMI', 'smoking_status' and 'stroke'

- *Libraries*

Pandas, Numpy, Seaborn, Matplotlib, Sklearn/Scikit-learn,Pickle, Joblib.

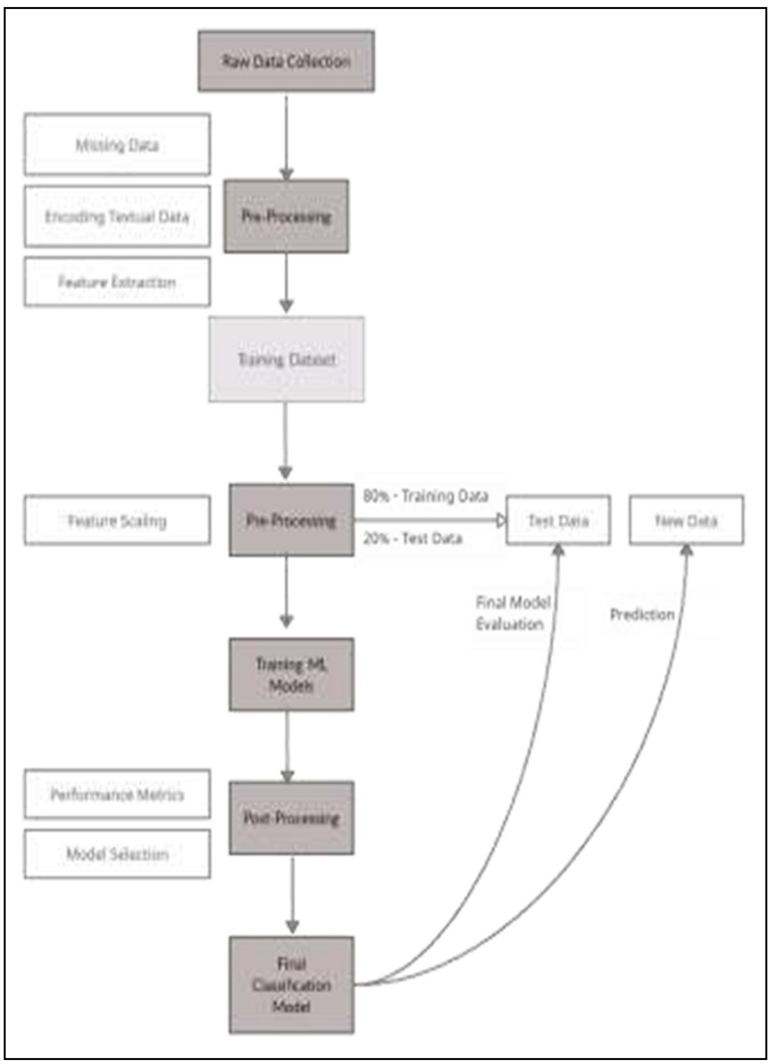


Fig.4 Workflow

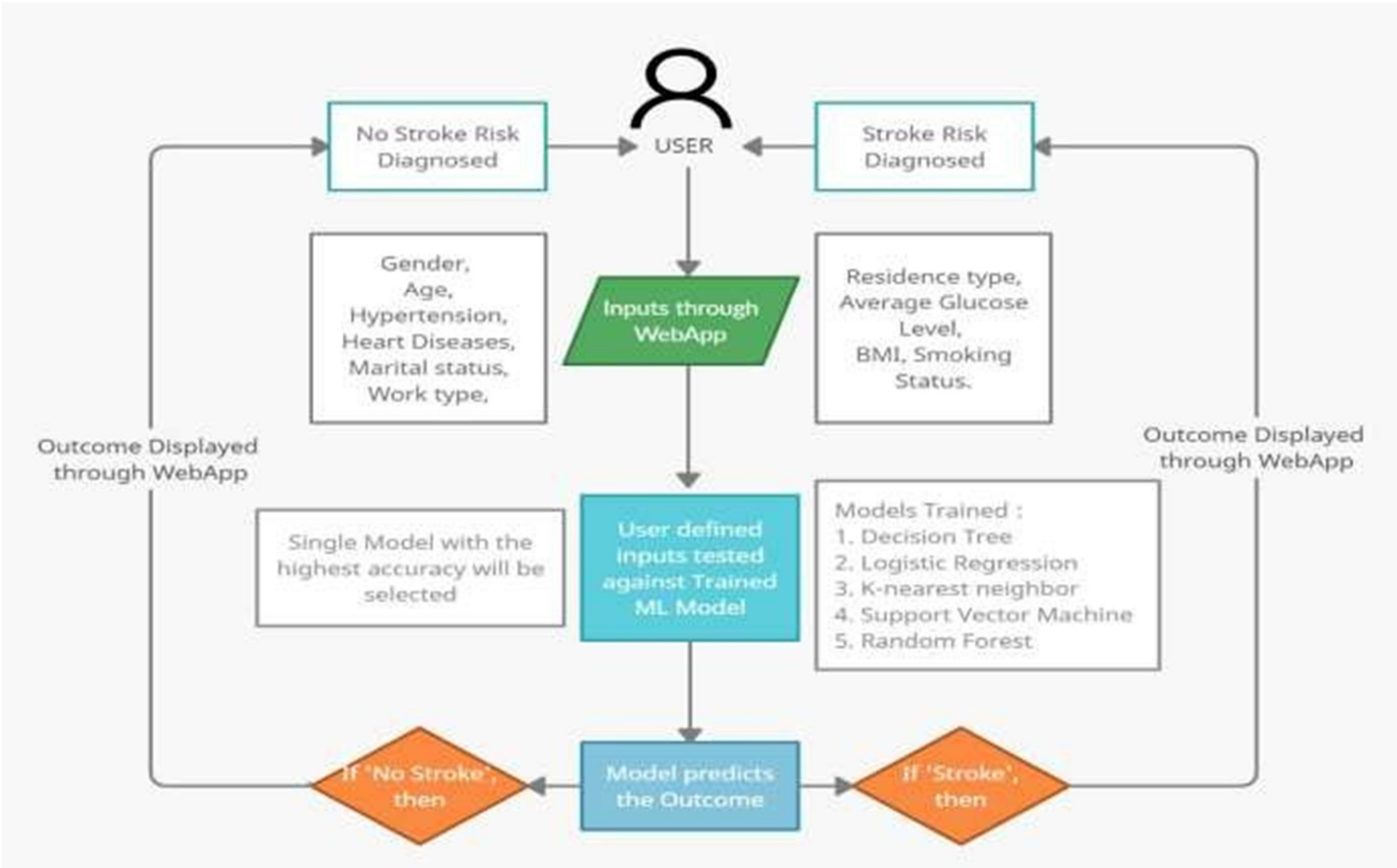


Fig.5 System Architecture

a) Clean the missing values in both training and testing data.
b) Applying Label Encoder to convert objects into integer.
c) Splitting the data in training and testing data.
d) Training ML Models:
- Decision Tree
- Logistic Regression
- KNN (K-nearest neighbor)
- SVM (Support Vector Machine)
- Random Forest

e) Calculating accuracy score for each model.
f) Model Selection with highest accuracy score
g) Create a GUI and extract that model into GUI module
h) Enter the new data for which stroke has to be predicted
i) Result: -Predicted data with respect to selected model.

The following Results have been generated:
- The lowest accuracy is given by Logistic RegressionAlgorithm i.e. (76.96%)
- The highest accuracy is given by Random ForestAlgorithm i.e. (98.56%)

Below is the Table of Evaluation parameters vs Models result (This was obtained from best dataset split for training and testing i.e. 70% for training and 30% for testing):

Evaluation parameters / Models	Accuracy Score	Recall Score	Precision Score	F1 Score
Decision Tree	97.39	100.00	94.95	97.41
Logistic Regression	76.96	82.22	73.87	77.82
KNN	91.64	100.00	85.86	92.16
SVM	82.00	88.84	77.73	82.92
Random Forest	**98.56**	**100.00**	**97.54**	**98.56**

Table 1 – Evaluation Parameters vs Models Result

So, we then imported the random forest trained model for testing against user-defined data and it is shown in Figure 3.

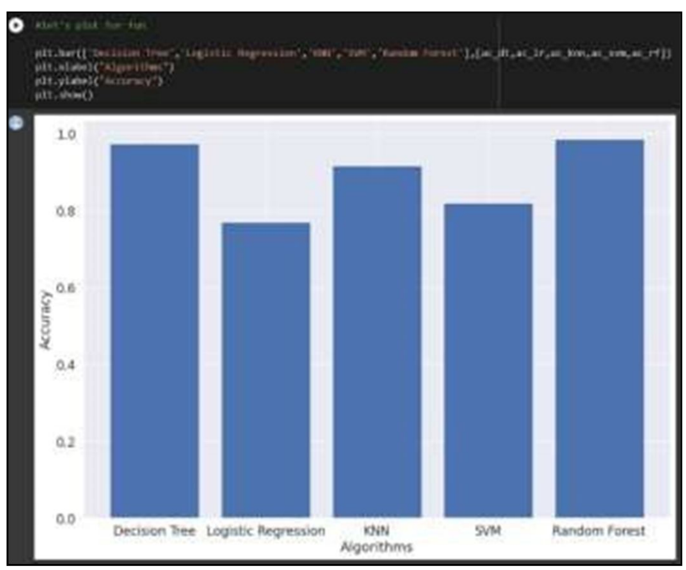


Figure 6 – Exporting Trained Model

The Comparison between all the Trained Models accuracy score is shown in the Figure 4 with the help of Matplotlib library of Python. Homepage will help user to input the necessary data required for stroke prediction linked and the GUI part is made quite flexible for common people. Knowledge of the desired output helps to reach the destination effectively.

As given in the Implementation, the system has been build using 5 different ML algorithms to get the best possible accuracy using our dataset

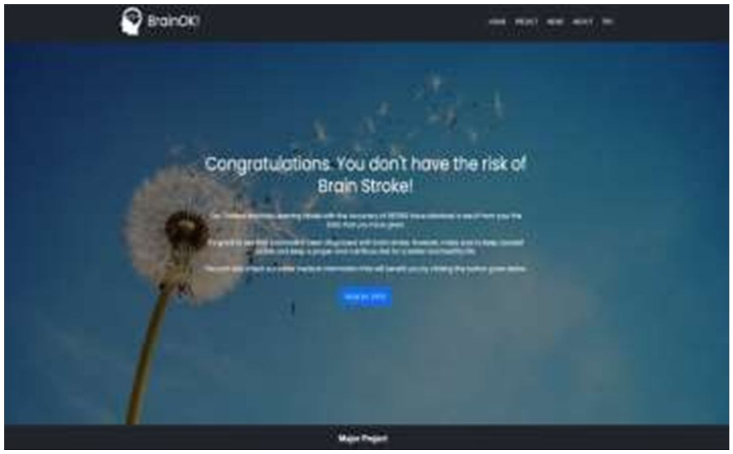


Fig.7 Trained Models Accuracy Score comparison

The proposed system helped us to analyse the best possible wayto take inputs from the user in the GUI implementation part of our project shown in the Figure 5 and with the data that is provided to the GUI for stroke risk prediction, the Random Forest model, which was trained with the dataset, was used andthe new data provided by the user was tested against the trainedmodel.

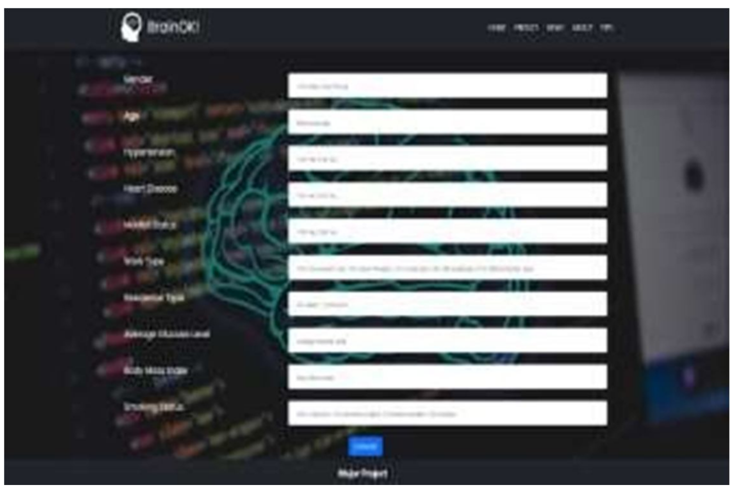


Figure 9

V. CONCLUSION

After the literature survey, we came to know various pros and cons of different research papers and thus, proposed a system that helps to predict brain strokes in a cost effective andefficient way by taking few inputs from the user side and predicting accurate results with the help of trained Machine Learning algorithms. Thus, the Brain Stroke Prediction systemhas been implemented using the given 5 Machine Learning algorithm given a highest accuracy of 98.56%. The system is therefore designed providing simple yet efficient User Interfacedesign with an empathetic approach towards their users and patients. 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.

VI. FUTURE SCOPE

1) Increasing the accuracy of the model.
2) Additional information about brain stroke can beexplained.
3) Allowing users to visualize their results based on theirinputs.

VII. ACKNOWLEDGMENT

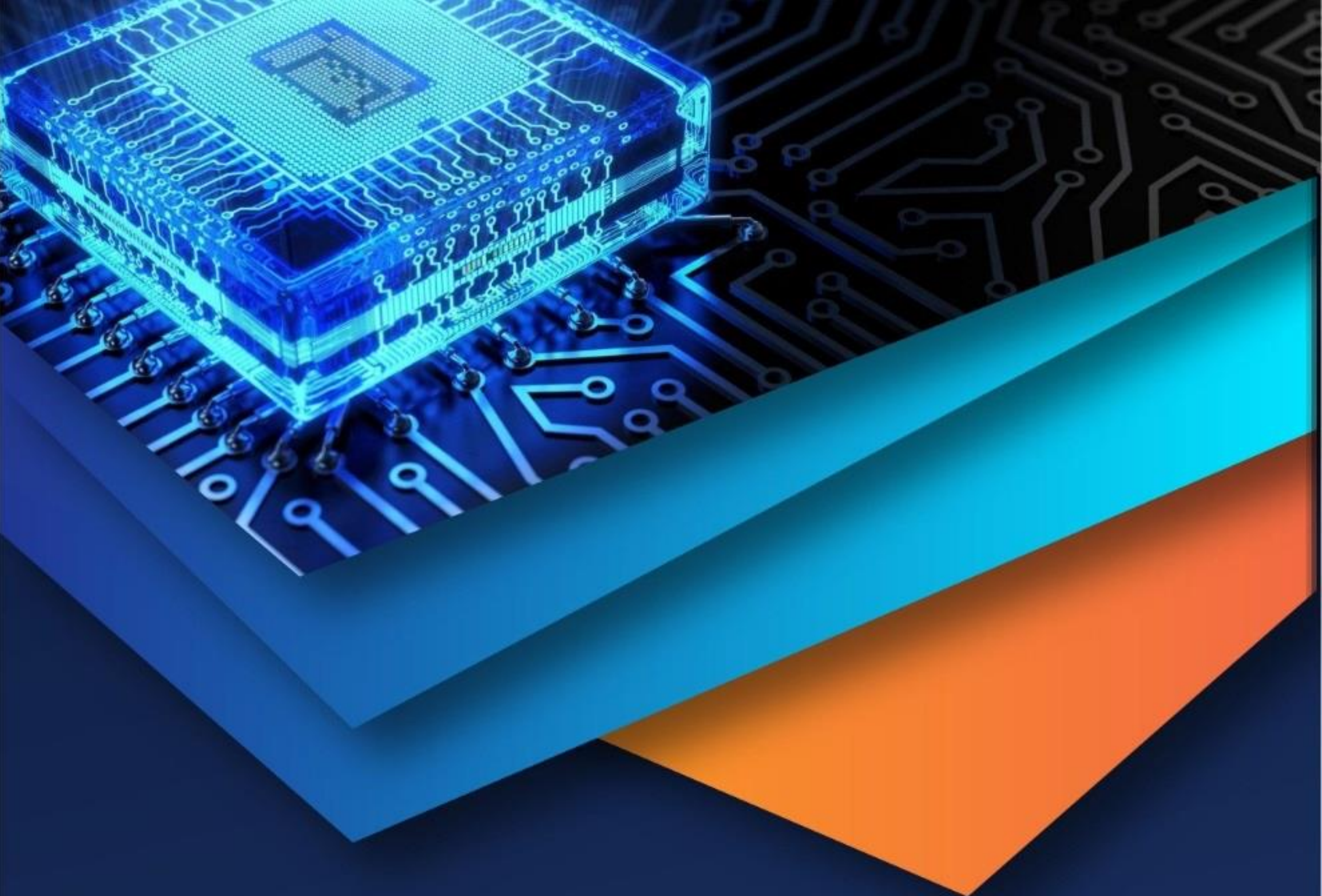
We are grateful to Prof. S. B. Ghawate for being our project mentor and assisting us in every step of the route.

Also, we would like to express our gratitude to H.O.D. Prof. S. N. Shelke for his unwavering encouragement and support during every phase of our project.

Lastly, we would like to thank all project stakeholders who were associated with the project and helped in its planning and execution. The project named “Sign Language Translation” would not have been possible without the extensive support of people who were directly or indirectly

REFERENCES

[1] Kalpesh Patil, Mandar Kulkarni, Anand Sriraman, and Shirish Karande. Deep learning-based car damage classification. pages 5054, 12 2017.
[2] World Health Organization, 2002. The world health report 2002: reducing risks, promoting healthy life. World Health Organization
[3] Pandian, J.D. and Sudhan, P., 2013. Stroke epidemiology and stroke care services in India. Journal of stroke, 15(3), p.128.
[4] Banerjee, T.K. and Das, S.K., 2016. Fifty years of stroke researches in India. Annals of Indian Academy of Neurology, 19(1), p.1.
[5] Ga, D., 2008. Fisher m, macleod m, Davis Sm. Stroke. Lancet, 371(9624), pp.1612- 23.
[6] Radic, B., 2017. Diagnosis and Treatment of Carotid Artery Stenosis. J. Neurol. Stroke, 7(3), pp.9-12.
[7] Ahirwar, D., Shakya, K., Banerjee, A., Khurana, D. and Roy, S., Simulation studies for non invasive classification of Ischemic and Hemorrhagic Stroke using Near In- frared Spectroscopy.
[8] A brain stroke detection model using soft voting based ensable using machine learning classifier , A.Shrinivas, Joseph Prakash Moseganti, year=2023, publisher: Elsevier
[9] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC97741



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY