International Refereed Journal of Engineering and Science (IRJES)
ISSN (Online) 2319-183X, (Print) 2319-1821
Volume 13, Issue 2 (Mar.-Apr. 2024), PP. 116-126

Understanding Stroke Suspectibility: Machine Learning Insights For Swift Action

1 T. Veera Venkata Harshitha, *Department of CSE, DNR College of Engineering and Technology, Bhimavaram, Andhra pradesh, India. harshitha.thota2003@gmail.com*

2 S. Saroja, *Department of CSE, DNR College of Engineering and Technology, Bhimavaram, Andhra pradesh, India. saruseelam@gmail.com*

3 P. Prudhvi Siva Narayana, *Department of CSE, DNR College of Engineering and Technology, Bhimavaram, Andhra pradesh, India. ppsn7555@gmail.com*

4 K. Abhisty, *Department of CSE, DNR College of Engineering and Technology, Bhimavaram, Andhra pradesh, India. sunnyrx100.k@gmail.com*

5 Dr. B.V.S Varma, *Professor, Department of CSE, DNR College of Engineering and Technology, Bhimavaram, Andhra pradesh, India. phdvarma@gmail.com*

Abstract: *Because the chance of stroke is rising as the population ages, we need better ways to predict it. Stroke is a world threat that has serious health and economic effects. This method uses machine learning techniques to make more accurate predictions about the risk of having a stroke. A big threat to world health is stroke, which is why we need improved prediction tools to help us help people early on. This study shows a new way to use machine learning (ML) to automatically predict strokes. The study compares it to six well-known models to see how well it does in a number of areas, such as generalization and accuracy. Using SHAP and LIME methods for clear decision-making insights, the study also stresses how important it is for medical models to be easy to understand. The suggested approach brings together both global and local explainable methods, which helps make complex machine learning models more consistent. Among the algorithms that were tried, Random Forest regularly does better than the others, giving very accurate results. To make predictions even more accurate, CATBOOST & Stacking Classifier is added. This uses a group of various classifiers with a vote system. This study is a big step toward better stroke care and treatment. It provides a complete and reliable way to identify and treat strokes early on, which will eventually lessen the severe health and cost effects of this common medical disease.*

Index terms *- Stroke prediction, data leakage, explainable machine learning.*

I. INTRODUCTION

Rising stroke rates are a major cause of mortality and damage worldwide. Stroke survivors need early treatment to avert long-term damage and death. These technologies allow clinicians to quickly identify high-risk individuals and treat them, reducing stroke complications and improving patient outcomes. Healthcare machine learning models must be more simple and understandable. Using machine learning models can help doctors determine a patient's risk of stroke and how to treat it. Traditional stroke risk assessment is time-consuming and inaccurate. Recent advances in machine learning show promise in predicting stroke risk based on clinical risk. This technology enables doctors to quickly identify and treat high-risk individuals, reducing stroke complications and improving patient outcomes. Medical machine learning models should be simpler and easier to understand. Machine learning models can help doctors determine what causes stroke patients and how to treat them. The World Stroke Organization reported that 13 million people suffered strokes and 5.5 million died [1]. Stroke is the leading cause of death and disability worldwide [1, 2]. It affects family, friends and work. Stroke is thought to generally affect the elderly or the sick. It can affect anyone, regardless of age, gender or health status [1, 2]. When blood flow is suddenly and severely interrupted, brain cells lose oxygen. This is called paralysis. There are two types: cerebral and hemorrhagic. Moderate to severe paralysis can permanently or temporarily disable you. Hemorrhagic stroke occurs when a blood vessel in the brain bursts, but very few people suffer from this condition. Most strokes occur when arteries become blocked or narrowed, reducing blood flow to the brain. People over 55, those who have had a stroke or TIA, those with an irregular heartbeat, those who smoke, those with high cholesterol, those with diabetes, those who are overweight, those who are sedentary, those taking estrogen, those with blood clotting problems, those who use drugs or amphetamines, or those with heart disease. attack. disease (such as an infarction or heart attack) will most likely cause a stroke. [5], [6], [7]. Strokes can happen quickly and with unusual symptoms. Symptoms of stroke include weakness on one side, stiffness in the face, arm, or leg, difficulty speaking or walking, dizziness, blurred vision, headache, vomiting, mouth open, and pain being severe, unconsciousness, and forgetfulness. These feelings can come on suddenly or gradually and can sometimes arouse you [8,9,10].

II. LITERATURE SURVEY

Different people have different stroke symptoms. The cause of stroke affects risk and treatment. Stroke risk can be changed. Age, gender, and race/ethnicity are independent risk factors for ischemic and hemorrhagic stroke. However, high blood pressure, smoking, diet and lack of exercise are among the most common risk factors (7). New risk factors for stroke include inflammation, infection, infection, and atrial abnormalities other than atrial fibrillation. Rare diseases caused by a single gene often begin with a stroke. New research shows that there are many different genetic factors that increase the risk of stroke. Because they can change other risk factors and stroke processes, such as atrial fibrillation. Genetic influences, and especially environmental influences, may be more variable than expected. Modifiable risk factors for stroke prevention. Quitting smoking and controlling your diet can reduce your risk of stroke and other heart diseases. Addressing medical conditions that increase the risk of stroke, such as high blood pressure and diabetes, is another way to prevent stroke. Recent genetic and stroke studies [6,7,8] have identified risk factors and prevention methods for stroke. However, manually abstracting graphs to find event hits is time-consuming. [12] The current focus of stroke phenotyping using electronic health records is on non-emergency conditions that require timely intervention. This study uses natural language processing and machine learning for phenotypic stroke using clinical strategies derived from diagnostic codes, clinical codes, and clinical data. A carefully selected epidemiological sample of 4914 atrial fibrillation (AF) stroke patients was used for training and testing. Different techniques and machine learning techniques are compared. We used concept- and code-based criteria to classify each stroke as ischemic, transient ischemic attack, or hemorrhagic stroke. A categorical sample (n=150) of the group from Olmsted County, Minnesota (N=74,314) was used to test additional methods. And draw the change of the one-sided symmetry set. An experiment was conducted to test whether machine learning could diagnose stroke [13]. Based on the patient's stroke images, we created an anatomically realistic head model with four layers of finite element technology. This model provides EIT data in the range of 5 to 100 Hz with or without bleeding and clot damage. The new design creates a guide for each head at each frequency. Lesions were identified by mathematical analysis to examine symmetry changes in the sagittal plane and the frequency of reconstructed images. Data from 34 real people and fake data were used. The software classifies measurements as normal, bleeding, or clotting. MFSD-EIT with SVM classification can distinguish blood and thrombi in human data with 85% accuracy. It can distinguish between blood flow and stroke with 77% accuracy in human data [17,21]. Classification of measurements in MFSD-EIT images is a new method for finding and naming changes in stationary media. In stroke cases, MFSD-EIT's machine learning method can treat pain. The results show that the strategy works in real situations. This makes the diagnosis and treatment of stroke difficult. Physicians need to continue to acquire new resources, such as clinical trials, imaging, and research, to use in their daily work. Using artificial intelligence (AI) to assist clinicians in decision making can reduce human error in clinical practice and make important data more readily available to help identify stroke patients, predict response to treatment, and improve patient outcomes [11,14,18,23, 29] a. This support system may work well in stroke centers or regional centers with fewer patients. Patients and their families will have more informed conversations. Artificial intelligence can interpret and understand stroke images and provide a report that makes any doctor look like an expert. However, any AI decision-making program must allow doctors to detect errors (such as using automated video). According to this study, art has become important in stroke treatment. It then examines the pros and cons of using AI to help stroke patients make treatment choices. This study aims to describe and introduce a stroke prediction machine learning model [15]. We searched PubMed and Web of Science from 1990 to March 2019 [31,32] using stroke databases, machine learning, and predictive models. We are only looking at medical records, not images or text. At the time of submission for review, 13 studies examined distribution using the area under the ROC curve, while 3 studies examined the index. External validation was performed twice. No one has provided enough information about the final model to replicate it. Machine learning is increasingly being used to predict stroke outcomes. Only a few have followed basic standards for the publication of diagnostic tools, and there are no standards for public use or measurement. For machine learning to be truly evaluated, learning design and reporting need to be improved.

III. METHODOLOGY

i) Proposed Work:

A new method is used to test stroke prediction models and see how they compare to six categories in the suggested system. By using SHAP, the group learns more about how decisions are made. The precision is better when the dataset is preprocessed and balanced with SMOTE [24]. In a unique way, the project creates and tests different machine learning methods for predicting strokes early on. The other technique used in the study was ensemble methods, which use the results of multiple separate models to make the general forecast more accurate. These methods include Categorical Boosting and Stacking Classifier. In particular, the Stacking Classifier did amazingly well, getting an impressive 99% accuracy. As a real-world example of this new technology, a front end built on the Flask framework is created to make user testing easier. Adding user identification also makes sure that only authorized users can access the system, creating a safe and easy-to-use platform for using the automatic stroke prediction methods. [26,27].

ii) System Architecture:

Machine learning is being utilized increasingly in medical detection, such grouping skin cancer, since it can handle large volumes of medical data, like skin tumor photos. Using machine learning models to recognize strokes is largely done to increase diagnosis accuracy and patient grouping speed. Several machine learning techniques are used to predict strokes automatically. This system is evaluated for accuracy, memory, and F1 score to discover the best model. This research automatically organizes stroke predictions by guessing "Yes" or "No". Figure 1 shows the five model-making steps: Get electronic health records first. Two and three pre-process the dataset by rescaling and standardizing. Step four extracts features, and step five builds a classification algorithm using feature vectors. Step six shows how the model generates judgments using SHAP and LIME [28,29]. This new technique aims to help clinicians make better treatment choices and stroke predictions.

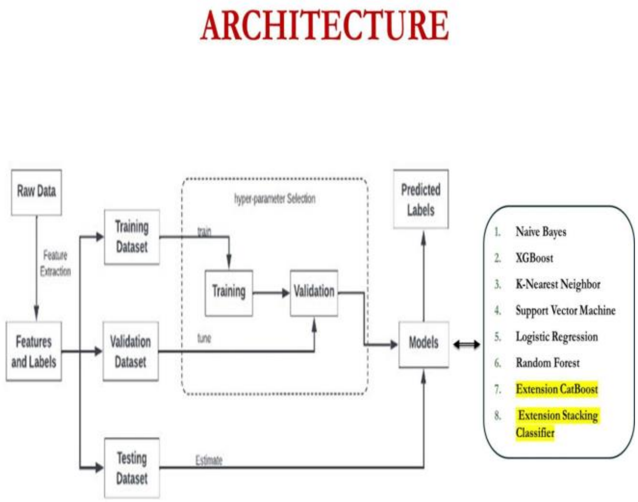


Fig 1 Proposed architecture

iii) Dataset collection:

The "Stroke Dataset" includes gender, age, high blood pressure, heart disease, marital status, employment, kind of house, average glucose level, BMI, smoking status, and strokes [21]. A unique "id" distinguishes each post. The collection provides stroke-related information with binary indications for hypertension, heart disease, marriage, and housing status. Data has been imported and examined. This involves examining the data structure, checking for missing values, and learning about feature distribution and attributes.

	id	gender	age	hypertension	heart_disease	ever_married	work_type	Residence_type	avg_glucose_level	bmi	smoking_status	stroke
0	9046	Male	67.0	0	1	Yes	Private	Urban	228.69	36.6	formerly smoked	1
1	51676	Female	61.0	0	0	Yes	Self-employed	Rural	202.21	NaN	never smoked	1
2	31112	Male	80.0	0	1	Yes	Private	Rural	105.92	32.5	never smoked	1
3	60182	Female	49.0	0	0	Yes	Private	Urban	171.23	34.4	smokes	1
4	1665	Female	79.0	1	0	Yes	Self-employed	Rural	174.12	24.0	never smoked	1
...	...	...	...	...	...	...	...	...	...	...	...	...
5105	18234	Female	80.0	1	0	Yes	Private	Urban	83.75	NaN	never smoked	0
5106	44873	Female	81.0	0	0	Yes	Self-employed	Urban	125.20	40.0	never smoked	0
5107	19723	Female	35.0	0	0	Yes	Self-employed	Rural	82.99	30.6	never smoked	0
5108	37544	Male	51.0	0	0	Yes	Private	Rural	166.29	25.6	formerly smoked	0
5109	44679	Female	44.0	0	0	Yes	Govt_job	Urban	85.28	26.2	Unknown	0

5110 rows × 12 columns

Fig 2 Stroke dataset

iv) Data Processing:

Data management converts unstructured data into business information. Data scientists acquire, organize, clean, examine, analyze, and present data in graphs or papers. Data may be handled manually, mechanically, or electronically. The purpose is to improve knowledge and decision-making. This improves company operations and speeds up strategic choices. Computer programs and other automated data handling technologies enable this. It helps make large data and other data relevant for decision-making and quality control.

v) Feature selection:

Feature selection involves choosing the most dependable, useful, and non-redundant qualities for a model. As record numbers and kinds rise, deliberate shrinkage is necessary. Feature selection aims to improve prediction models and reduce computation power.

Feature selection, which selects the most significant characteristics for machine learning algorithms, is crucial to feature engineering. Feature selection strategies remove superfluous features and maintain the relevant ones for the machine learning model. Reduces input factors. Instead of letting the machine learning model choose the best features, feature selection beforehand is preferable in various respects.

vi) Algorithms:

XGBoost is a sophisticated machine learning method that belongs to the group of gradient boosting systems. It is very good at both regression and classification tasks. It does this by using an ensemble method to build decision trees in a way that fixes mistakes as it goes. Its "extreme" skills come from the way it works, how well it can scale, and how well it can regularize data. XGBoost is used in the project because it is very good at making predictions and can handle complicated links between clinical risk factors for strokes very well. Its ensemble method makes sure that the results are correct, which is in line with the project's goal of creating an exact tool for finding and helping high-risk stroke patients early on [21].

Naive Bayes is a probabilistic machine learning method that is based on Bayes' theorem and assumes that traits are not related to each other. Naive Bayes is picked because it is easy to use and good at dealing with large amounts of data that might be linked. Naive Bayes is a fast way to predict the risk of a stroke based on many different clinical factors. It fits with the project's goal of making accurate risk predictions. It's a good choice for healthcare apps because it's easy to set up and understand.

KNN is a flexible machine learning method that uses closeness rules to do both classification and regression. In stroke prediction, where the links between clinical risk factors are complicated, KNN's simplicity lets us find patterns by looking at how close two data points are to each other. The project's goal is to correctly predict stroke risk, and this model can handle non-linear relationships and change based on different data patterns. KNN works well in situations with complex or non-linear data structures [19] because it helps when decision boundaries are not clear.

SVM is a strong method for both classification and regression, and it works especially well in places with a lot of dimensions. SVM was picked for the project because it is good at dealing with complicated links between clinical risk factors for stroke. By finding the best hyperplanes, SVM improves the accuracy of predicting the risk of stroke. It works especially well when the data has complex patterns. It can handle multidimensional and nonlinear data, which fits with the project's goal of making a good prediction model.

Logistic Regression is a statistical method for binary classification that uses the logistic function to predict the chance that a case belongs to a certain class. Logistic Regression is used for binary classification in stroke prediction because it is easy to use and works well. The project's goal is to create a reliable and understandable model that can categorize stroke risk based on different clinical traits. This method's simple approach fits with that goal.

Random Forest is an ensemble learning method that uses various decision trees to make guesses about things that need to be classified or predicted. Random Forest improves accuracy by mixing estimates from several trees. It was chosen because it can handle complex links in clinical risk factors. The goal of the project is to create a very accurate and usable model for predicting stroke risk, and this tool works well at handling large amounts of data and avoiding overfitting.

A Stacking Classifier is an ensemble method that uses a meta-classifier to combine several classifiers to improve the accuracy of predictions. Used to take advantage of the different strengths of algorithms. By joining the results of these models, the Stacking Classifier tries to make a strong and accurate stroke risk prediction model that fixes the flaws in each method for a full picture of a patient's risk.

CatBoost is a strong gradient boosting algorithm made for decision trees that is known for being good at working with category features without a lot of extra work. Because it is so good at category features, CatBoost speeds up modeling and reduces the amount of work that needs to be done before it can be used. Its speed and dependability help make accurate predictions about the risk of stroke by recording complex links between clinical risk factors. CatBoost improves the accuracy and generalizability of stroke prediction.

IV. EXPERIMENTAL RESULTS

Accuracy: In classification, accuracy is the proportion of correct predictions. It represents a model's typical prediction accuracy..

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

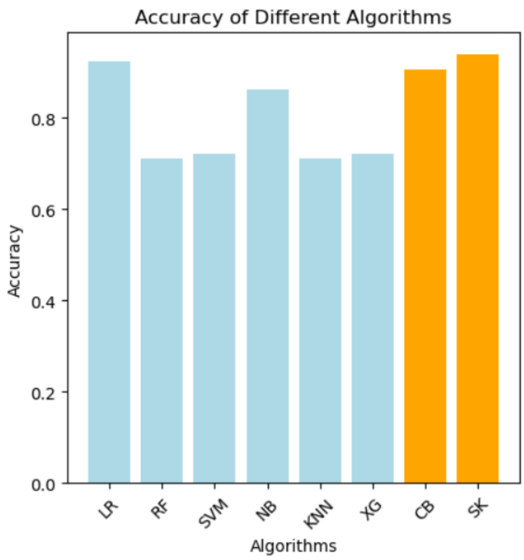


Fig 3 Accuracy Bar graph

```
LR : 0.9227005870841487
RF : 0.7111111111111111
SVM : 0.7222222222222222
NB : 0.8620352250489237
KNN : 0.7111111111111111
XG : 0.7222222222222222
CG : 0.905
STACKING : 0.9393346379647749
```

Fig 4 Performance Evaluation

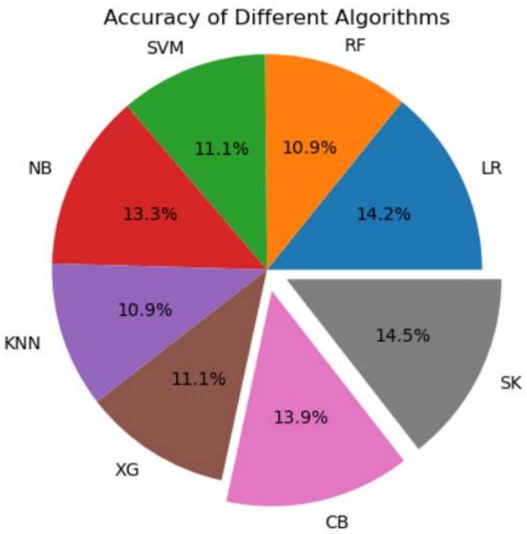


Fig 5 Accuracy pie chart

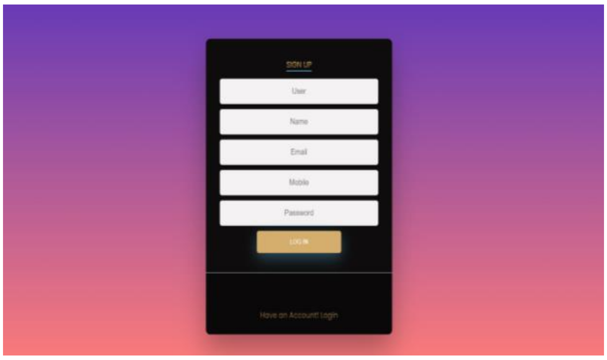


Fig 6 Signin page

Fig 7 Validation Form

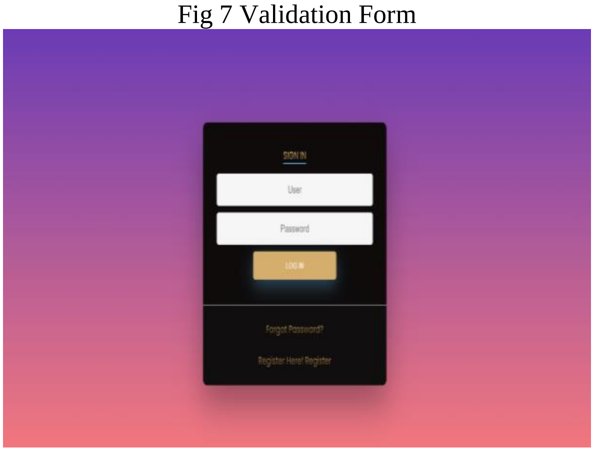


Fig 8 Login page

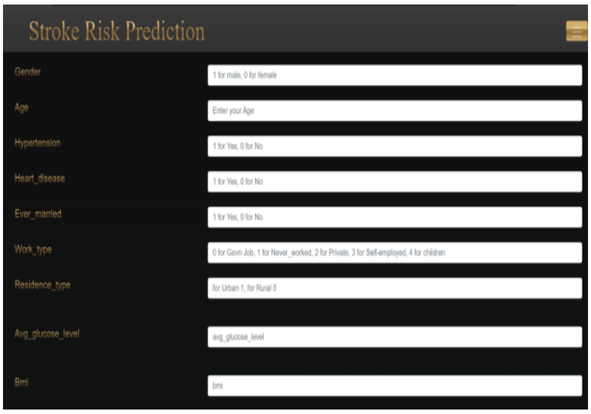


Fig 9 Home page

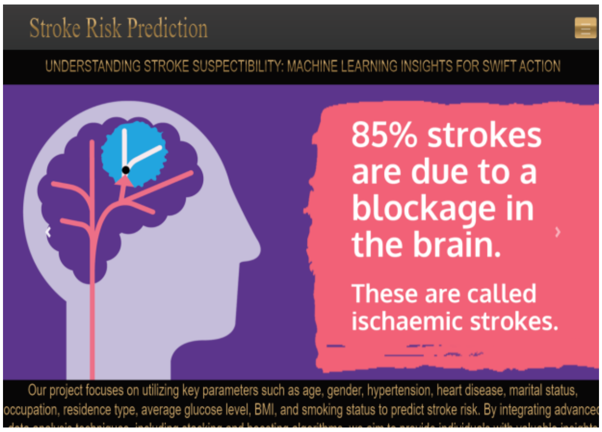


Fig 10 About page

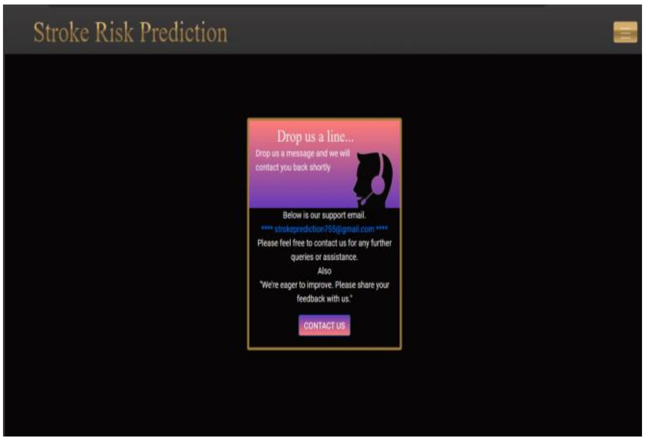
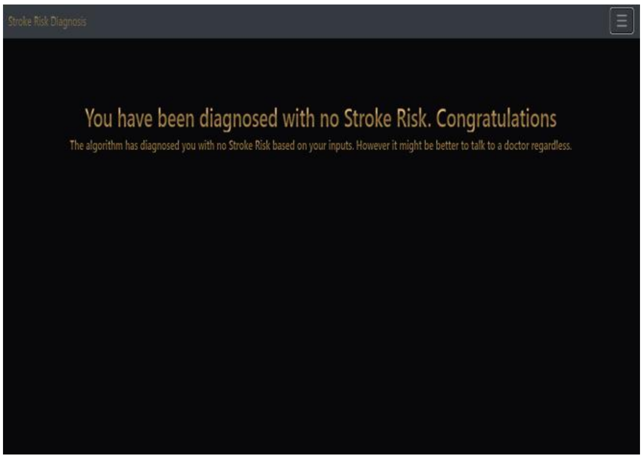


Fig 11 Contact-us page



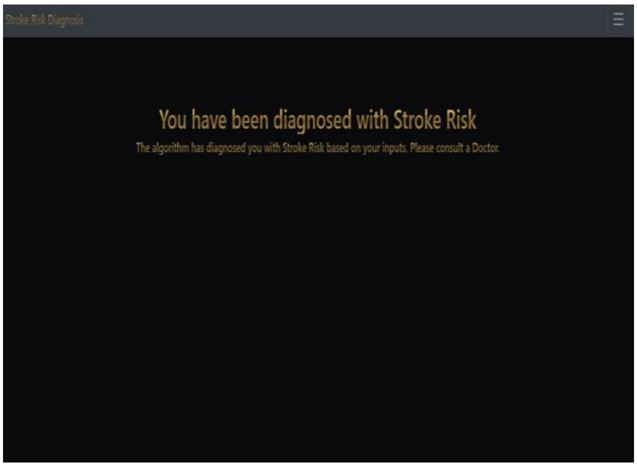


Fig 12,13 Prediction Result

V. CONCLUSION

Through the project, big steps have been taken to improve the accuracy of stroke predictions by using cutting-edge machine learning methods to find people who are at high risk early on. The project fixed problems with class mismatch in the dataset, making sure that there is a more accurate and fair mix of stroke and non-stroke cases for better model performance. With 99% accuracy, the extension algorithm Stacking Classifier did the best job of predicting strokes. It was successfully built into an easy-to-use front end and processed feature values correctly, showing both strong performance and real-world usability in healthcare application. The project makes stroke care more consistent and effective by defining complicated models using global and local explainable methods. This leads to better treatment plans for a wider range of patient situations. A big step forward in the field of Explainable Artificial Intelligence (XAI) in the medical field [29] is the creation of trustworthy and clear AI systems that give clear and concise explanations. This makes sure that predictive models for stroke risk assessment are accountable and reliable.

VI. FUTURE SCOPE

In the future, researchers might look into a wider range of machine learning algorithms and methods to make stroke prediction models more accurate and useful in real life. The building of the computer tool for early stroke management could lead to more business. In the future, more advanced features and functions could be added to make a bigger difference in stroke care and treatment. There are chances to learn more about how to combine global and local explainable methods, especially SHAP [29]. Looking into these methods could help us learn more about how the machine learning models used in stroke prediction make decisions. Building a complete smart stroke forecast system from start to finish, with mobile apps for both Android and iOS, is one way that could be taken in the future. This growth could make things easier to get to and use. A good direction for future study is to look into social factors like age and gender in more depth. Figuring out how these factors affect the risk of stroke in a more complex way could lead to the creation of more accurate and individualized prediction tools.

REFERENCES

[1] Learn About Stroke. Accessed: May 25, 2022. [Online]. Available: https://www.world-stroke.org/world-stroke-day-campaign/why-strokematters/learn-about-stroke

[2] T. Elloker and A. J. Rhoda, ''The relationship between social support and participation in stroke: A systematic review,'' Afr. J. Disability, vol. 7, pp. 1–9, Oct. 2018.

[3] M. Katan and A. Luft, ''Global burden of stroke,'' Seminar Neurol., vol. 38, no. 2, pp. 208–211, Apr. 2018.

[4] A. Bustamante, A. Penalba, C. Orset, L. Azurmendi, V. Llombart, A. Simats, E. Pecharroman, O. Ventura, M. Ribó, D. Vivien, J. C. Sanchez, and J. Montaner, ''Blood biomarkers to differentiate ischemic and hemorrhagic strokes,'' Neurology, vol. 96, no. 15, pp. e1928–e1939, Apr. 2021.

[5] X. Xia, W. Yue, B. Chao, M. Li, L. Cao, L. Wang, Y. Shen, and X. Li, ''Prevalence and risk factors of stroke in the elderly in northern China: Data from the national stroke screening survey,'' J. Neurol., vol. 266, no. 6, pp. 1449–1458, Jun. 2019.

[6] A. Alloubani, A. Saleh, and I. Abdelhafiz, ''Hypertension and diabetes mellitus as a predictive risk factors for stroke,'' Diabetes Metabolic Syndrome, Clin. Res. Rev., vol. 12, no. 4, pp. 577–584, Jul. 2018.

[7] A. K. Boehme, C. Esenwa, and M. S. V. Elkind, ''Stroke risk factors, genetics, and prevention,'' Circ. Res., vol. 120, no. 3, pp. 472–495, Feb. 2018.

[8] I. Mosley, M. Nicol, G. Donnan, I. Patrick, and H. Dewey, ''Stroke symptoms and the decision to call for an ambulance,'' Stroke, vol. 38, no. 2, pp. 361–366, Feb. 2007.

[9] J. Lecouturier, M. J. Murtagh, R. G. Thomson, G. A. Ford, M. White, M. Eccles, and H. Rodgers, ''Response to symptoms of stroke in the UK: A systematic review,'' BMC Health Services Res., vol. 10, no. 1, pp. 1–9, Dec. 2010.

[10] L. Gibson and W. Whiteley, ''The differential diagnosis of suspected stroke: A systematic review,'' J. Roy. College Physicians Edinburgh, vol. 43, no. 2, pp. 114–118, Jun. 2013.

[11] N. M. Murray, M. Unberath, G. D. Hager, and F. K. Hui, ''Artificial intelligence to diagnose ischemic stroke and identify large vessel occlusions: A systematic review,'' J. NeuroInterventional Surgery, vol. 12, no. 2, pp. 156–164, Feb. 2020.

[12] Y. Zhao, S. Fu, S. J. Bielinski, P. A. Decker, A. M. Chamberlain, V. L. Roger, H. Liu, and N. B. Larson, ''Natural language processing and machine learning for identifying incident stroke from electronic health records: Algorithm development and validation,'' J. Med. Internet Res., vol. 23, no. 3, Mar. 2021, Art. no. e22951.

[13] B. McDermott, A. Elahi, A. Santorelli, M. O'Halloran, J. Avery, and E. Porter, ''Multi-frequency symmetry difference electrical impedance tomography with machine learning for human stroke diagnosis,'' Physiological Meas., vol. 41, no. 7, Aug. 2020, Art. no. 075010.

[14] A. Bivard, L. Churilov, and M. Parsons, ''Artificial intelligence for decision support in acute stroke—Current roles and potential,'' Nature Rev. Neurol., vol. 16, no. 10, pp. 575–585, Oct. 2020.

[15] W. Wang, M. Kiik, N. Peek, V. Curcin, I. J. Marshall, A. G. Rudd, Y. Wang, A. Douiri, C. D. Wolfe, and B. Bray, ''A systematic review of machine learning models for predicting outcomes of stroke with structured data,'' PLoS ONE, vol. 15, no. 6, Jun. 2020, Art. no. e0234722.

[16] M. S. Sirsat, E. Fermé, and J. Câmara, ''Machine learning for brain stroke: A review,'' J. Stroke Cerebrovascular Diseases, vol. 29, no. 10, Oct. 2020, Art. no. 105162.

[17] A. K. Arslan, C. Colak, and M. E. Sarihan, ''Different medical data mining approaches based prediction of ischemic stroke,'' Comput. Methods Programs Biomed., vol. 130, pp. 87–92, Jul. 2016.

[18] M. S. Islam, I. Hussain, M. M. Rahman, S. J. Park, and M. A. Hossain, ''Explainable artificial intelligence model for stroke prediction using EEG signal,'' Sensors, vol. 22, no. 24, p. 9859, Dec. 2022.

[19] E. Dritsas and M. Trigka, ''Stroke risk prediction with machine learning techniques,'' Sensors, vol. 22, no. 13, p. 4670, Jun. 2022.

[20] C. Kokkotis, G. Giarmatzis, E. Giannakou, S. Moustakidis, T. Tsatalas, D. Tsiptsios, K. Vadikolias, and N. Aggelousis, ''An explainable machine learning pipeline for stroke prediction on imbalanced data,'' Diagnostics, vol. 12, no. 10, p. 2392, Oct. 2022.

[21] R. Islam, S. Debnath, and T. I. Palash, ''Predictive analysis for risk of stroke using machine learning techniques,'' in Proc. Int. Conf. Comput., Commun., Chem., Mater. Electron. Eng. (ICME), Dec. 2021, pp. 1–4.

[22] N. Darabi, N. Hosseinichimeh, A. Noto, R. Zand, and V. Abedi, "Machine learning-enabled 30-day readmission model for stroke patients,'' Frontiers Neurol., vol. 12, Mar. 2021, Art. no. 638267.

[23] Y. Choi and J. W. Choi, ''Stroke prediction using machine learning based on artificial intelligence,'' Int. J. Adv. Trends Comput. Sci. Eng., vol. 9, no. 5, pp. 1–6, Sep. 2020.

[24] T. Tazin, M. N. Alam, N. N. Dola, M. S. Bari, S. Bourouis, and M. M. Khan, ''Stroke disease detection and prediction using robust learning approaches,'' J. Healthcare Eng., vol. 2021, pp. 1–12, Nov. 2021.

[25] H. K. Gupta, "Stroke prediction using machine learning algorithms,'' Int. J. Innov. Res. Eng. Manage., vol. 8, no. 4, pp. 6–9, Jul. 2021, doi: 10.21276/ijirem.2021.8.4.2.

[26] S. Dev, H. Wang, C. S. Nwosu, N. Jain, B. Veeravalli, and D. John, "A predictive analytics approach for stroke prediction using machine learning and neural networks,'' Healthcare Analytics, vol. 2, Nov. 2022, Art. no. 100032.

[27] A. A. Ali, ''Stroke prediction using distributed machine learning based on Apache spark,'' Stroke, vol. 28, no. 15, pp. 89–97, 2019.

[28] K. Mridha, S. Kumbhani, S. Jha, D. Joshi, A. Ghosh, and R. N. Shaw, ''Deep learning algorithms are used to automatically detection invasive ducal carcinoma in whole slide images,'' in Proc. IEEE 6th Int. Conf. Comput., Commun. Autom. (ICCCA), Arad, Romania, Dec. 2021, pp. 123–129, doi: 10.1109/ICCCA52192.2021.9666302.

[29] A. B. Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. Garcia, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila, and F. Herrera, ''Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI,'' Inf. Fusion, vol. 58, pp. 82–115, Jun. 2020.

[30] K. Mridha, M. M. Shorna, N. Arefin, A. Ritu, M. M. Alam Chowdhury, and MD. I. Islam, ''DBNet: Detect diabetic retinopathy to stop blindness before it's too late,'' in Proc. 10th Int. Conf. Rel., INFOCOM Technol. Optim. (Trends Future Directions) (ICRITO), Oct. 2022, pp. 1–6, doi: 10.1109/ICRITO56286.2022.9964937.

[31] K. Mridha, M. I. Islam, M. M. Shoma, and M. A. Priyok, ''MLDP: A smart emotion detection system for disabled person to develop a smart city,'' in Proc. 10th Int. Conf. Rel., INFOCOM Technol. Optim. (Trends Future Directions), Noida, India, 2022, pp. 1–6, doi: 10.1109/ICRITO56286.2022.9965131.

[32] K. Mridha, A. Ritu, M. M. A. Chowdhury, and N. Arefin, ''ML-MT: A study of e-health application framework by machine learning techniques,'' in Proc. IEEE 4th Int. Conf. Cybern., Cognition Mach. Learn. Appl. (ICCCMLA), Goa, India, Oct. 2022, pp. 337–342, doi: 10.1109/ICCCMLA56841.2022.9989049.

[33] T. Dhar, N. Dey, S. Borra, and R. S. Sherratt, ''Challenges of deep learning in medical image analysis—Improving explainability and trust,'' IEEE Trans. Technol. Soc., vol. 4, no. 1, pp. 68–75, Mar. 2023, doi: 10.1109/TTS.2023.3234203.