An Enhanced Machine Learning Framework for Brain Stroke Risk Prediction Using Multi-Dataset Validation, Latent-Space Oversampling, and Clinical Feature Engineering

Anita Sahu[1], **Prof Sudeep Kishore Sharma**[2], **Prof. Prateek Gupta**[3]

[1]Research Scholar Department of CSE, Shriram Institute of Science & Technology, Jabalpur, M.P.
[2]Prof., Department of CSE, Shriram Institute of Science & Technology, Jabalpur, M.P.
[3]Prof., Department of CSE, Shriram Institute of Science & Technology, Jabalpur, M.P.

***

Abstract *- This paper presents a systematic comparative study of machine learning classification models for brain stroke symptoms detection using three independent datasets of varying scale, feature richness, and origin. This protocol successfully identifies one of the three candidate datasets (Strokesdataset.csv, 12,000 records, 16 features) as synthetically generated with randomly assigned stroke labels, uniform stroke rates across all clinical subgroups (~50%), and all three trained classifiers scoring AUC ≈ 0.50.*

Categorical features are encoded using Label Encoding for binary attributes and One-Hot Encoding for multi-class unordered attributes. Six domain-informed interaction and polynomial features are engineered from the base feature set — including age^2, age×glucose, age×hypertension, and a composite clinical risk score — expanding the feature matrix from 11 to 20 dimensions. Synthetic Three classification algorithms are trained and evaluated: Logistic Regression with balanced class weights (interpretable baseline), Random Forest with 50 estimators and depth constraint of 8, and XGBoost with scale_pos_weight calibrated to the empirical imbalance ratio. Models are evaluated using AUC-ROC as the primary metric, supplemented by F1score, precision, and recall for the stroke-positive class, with explicit threshold optimization using the Precision-Recall curve. This work contributes a validated end-to-end stroke detection pipeline, a reusable dataset quality assessment protocol,

Key Words: Brain Stroke Detection, Machine Learning, Classification Algorithms SMOTE, Class Imbalance, Feature Engineering, AUC-ROC, Random Forest, Comparative Study, Synthetic Data Detection.

1. INTRODUCTION

The incidence of stroke has been increasing globally, and it is now considered one of the leading causes of death and dis ability. Early intervention is crucial in preventing long-term disability and mortality associated with stroke. Traditional methods of predicting stroke risk, however, are often time-consuming and prone to errors. Recently, machine learning algorithms have shown great promise in accurately predicting stroke risk based on various clinical risk factors. By leveraging these algorithms, clinicians can identify high-risk patients and intervene early, potentially reducing the number of stroke-related complications and improving patient outcomes. Additionally, there is a growing need for transparency and explainability in machine learning models in healthcare. The use of an interpretable machine learning model can provide clinicians with valuable insights into the factors that contribute to a patient's stroke risk, thereby aiding in treatment decisions. The World Stroke Organization estimates that 13 million people worldwide experience a stroke each year, leading to 5.5 million fatalities [1]. Stroke affects all aspects of a patient's life, including their family, social environment, and work, and is one of the top causes of mortality and disability in the world [2]. In reality, anybody can be impacted, regardless of age, gender, or physical health [2]. A stroke is a rapid, serious disruption in blood flow to the brain that deprives brain cells of oxygen. It comes in ischemic and hemorrhagic varieties. Moderate to severe strokes can cause permanent or temporary damage, depending on their severity. Hemorrhagic strokes are uncommon; however, they are brought on by the rupture of a blood vessel in the brain. The most common type of stroke happens when an artery is blocked or narrows, preventing blood flow to the brain [3], [4]. Age over 55, prior stroke or TIA, arrhythmia, high blood pressure, carotid stenosis from atherosclerosis, smoking, high blood cholesterol, diabetes, obesity, inactivity, estrogen therapy, blood clotting disorders, cocaine or amphetamine use, and heart issues like infarction and cardiac arrest are all risk factors for stroke [5], [6], [7]. Strokes can occur suddenly, and their symptoms might vary and be unanticipated. The main symptoms of a stroke include paralysis on one side of the body, numbness in the face, arms, or legs, difficulty speaking or walking, dizziness, blurred vision, headache.

These sensations may come on suddenly or gradually, and in certain rare cases, they may cause you to become aware [8]. Stroke can impact both men and women, lowering their quality of life and putting a load on public health resources. The scientific community prioritizes building models for predicting strokes to avoid them, and AI plays a critical role in this endeavor because it is extensively employed for disease prevention.

1.1 Methods for Stroke Risk Prediction

Stroke is considered one of the most serious diseases of today's life. It can cause physical and mental deaths such as hemiplegia, speech disorder (aphasia), ataxia, blindness, forgetfulness, and dementia. According to the 2019 death index published by the world health organization (who) in December 2020, the top 10 causes of death accounted for 55% of the total number of deaths in 2019. Among them, 6 million people died of cerebrovascular disease, which was reported to be the second leading cause of death [9]. The United Nations states that if the population aged 65 and over reaches 7 percent or more in a country, that country is classified as elderly; an old life; when the rate is more than 20%, it means that it has entered super old life [10]. For this reason, the social problems of the elderly society have begun to gain importance and the elderly can be defined as segmental. In addition, the aging report of the international credit rating company Moody’s shows that countries such as Japan, Germany, and Italy have entered the elderly population since 2013, with the elderly accounting for more than 20%. It is reported that as many as 34 countries will enter the elderly age by 2030 [33]. The prognosis and health of patients after a stroke vary greatly depending on the age and location of the stroke. According to previous studies on stroke, more than 66% of all strokes occur in adults aged 65 and over [11]. In addition to these social problems, stroke morbidity and mortality should also become important social and economic problems. [12]. The main techniques used for neurological diagnosis in stroke diagnosis are brain mri and ct, and other studies have reported that biological diseases such as brain weakness, muscles, and electrocardiograms can also be used to diagnose and prevent stroke [13]. In addition, ultrasound, echocardiography, cerebral angiography, and single photon emission computed tomography (spect) have also been used to determine the cause of stroke [14]. In recent years, the widespread use of imaging such as ct and mri in stroke electrical diagnosis and claustrophobia in closed spaces have made testing and testing difficult. Place. There may also be errors in the test, so it is important to make decisions based on the doctor's clinical experience and empirical evidence. Second, the national institutes of health stroke scale (nihss), published by the national institutes of health, were used as a study to prevent stroke recurrence and assess early disability in stroke patients [15]. Although the nihss is a method to assess initial damage in stroke patients, there are limitations in the immediate detection of damage and problems in clinical and psychological studies. Recent studies have shown that many studies have used electrocardiography to predict and prevent stroke from atrial fibrillation, one of the causes of stroke [16], [17]. Atrial fibrillation (af), often associated with hypertension, is a major risk factor for stroke and is known to increase the risk of cerebral infarction by more than fivefold [18]. A previous study used clinical trials to identify stroke risk factors, including high blood pressure, smoking, obesity, and diabetes mellitus [19]. Therefore, there is a need for a method that will enable the elderly to evaluate the affected stroke patient and detect the possible consequences of the disease early and in a short time. To overcome these limitations, recent studies have tried to use statistics or machine learning to predict stroke, including certain risk factors. However, these methods have a limitation that they still give black results for predictions, which makes them difficult to interpret. According to the decision tree method, part of the analysis method can be used with the heuristic method, but there is now a need for research on predictive methods that can cover previous studies and also cover the consequences of disease development and translation. Chads2 is the score proposed by [20]. The name chads2 is an abbreviation for the method; the letter "c" indicates the details of patients with heart failure, the letter "h" indicates the details of the study of patients with heart failure, and the letter "a" indicates hypertensive patients aged 75 years and over. , "D" for diabetic patients, and finally "s2" for patients who have had a stroke, ischemic attack or thromboembolism in the past, a score of 2 is given. The chads2 score is widely used in clinical practice to predict stroke outcomes because it is easy to develop and evaluate. The disadvantage of this method is that it uses only five variables in decision making. [21] They developed an interpretable model using the Bayesian rule list (brl) to predict the one-year risk of stroke in patients diagnosed with atrial fibrillation. This method uses a decision list consisting of "if" words and then uses a structure called the Bayesian rule list, which is therefore a later part of the decision list. This thought may cause stroke.

1.2 Objectives

1. **To perform systematic data preprocessing and feature engineering**

- Handle missing values, outliers, and categorical variables.

- Apply normalization and scaling techniques to enhance data consistency.
- Implement feature selection and extraction for improved model efficiency.

2. To train and evaluate multiple Machine Learning models

- Implement algorithms such as Logistic Regression, Random Forest, XGBoost, AdaBoost, LightGBM, and CatBoost.
- Optimize each model using cross-validation and evaluate on standard metrics.

3. To design and train Deep Learning architectures for stroke prediction

- Develop and compare models such as Multi-Layer Perceptron (MLP), Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), and TabNet.
- Employ appropriate activation functions, optimizers, and regularization techniques.

4. To assess model performance using comprehensive evaluation metrics

- Measure Accuracy, Precision, Recall, F1-Score, ROC-AUC, and Confusion Matrix for each model.
- Generate classification reports and visualize performance through comparison plots.

5. To conduct a comparative analysis between ML and DL models

- Identify the best-performing models in terms of predictive power, generalization, and interpretability.
- Highlight the trade-offs between complexity, accuracy, and clinical applicability.
- To design and implement a reproducible end-to-end machine learning pipeline for brain stroke risk detection from structured patient data.
- To propose and validate a statistical protocol for assessing the quality and authenticity of medical datasets prior to modelling.
- To address severe class imbalance (52:1 ratio) using SMOTE without introducing data leakage into model evaluation.
- To engineer clinically meaningful interaction and polynomial features that capture non-linear stroke risk relationships beyond the base feature set.
- To train and systematically compare three classification algorithms — Logistic Regression, Random Forest, and XGBoost — under identical experimental conditions
- To perform decision threshold optimisation using the Precision-Recall curve to identify clinically appropriate operating points.
- To validate model generalisability by replicating the complete pipeline on an independent stroke dataset without modification.

2. LITERATURE SURVEY

This study [22] explores the impact of leveraging retinal images and clinical data for stroke detection and risk prediction. We propose a multimodal deep neural network that processes Optical Coherence Tomography (OCT) and infrared reflectance retinal scans, combined with clinical data, such as demographics, vital signs, and diagnosis codes. Our empirical findings establish the predictive ability of the considered modalities in detecting lasting effects in the retina associated with acute stroke and forecasting future risk within a specific time horizon. The experimental results demonstrate the effectiveness of our proposed framework by achieving 5% AUROC improvement as compared to the unimodal image-only baseline, and 8% improvement compared to an existing state-of-the-art foundation model. In conclusion, our study highlights the potential of retinal imaging in identifying high-risk patients and improving long-term outcomes.

The proposed [23] platform leverages a novel combination of SMOTE oversampling and logistic regression to address class imbalance in patient health records, improving the detection of stroke risk factors over existing methods. Our best model (logistic regression with SMOTE and standard scaling) achieves 93.2% accuracy with a substantially higher F1-score for the stroke-positive class than other models, indicating improved sensitivity to stroke cases. To bridge the gap between complex predictive models and end-users, we deploy this model in an intuitive web interface (built with Streamlit) that non-technical individuals and healthcare providers can easily use. By offering a practical and transparent tool for stroke risk screening, our work advances health informatics with an emphasis on accessibility, interpretability, and early intervention. Potential applications range from personal health self-assessment to integration in clinical workflows for preventive care, ultimately aiming to improve public health outcomes through early detection and intervention in stroke.

Authors in paper [24] proposed ML technique was explored in terms of metrics relating to both generalization capability and prediction accuracy. To give insight into the black-box machine learning models, we also studied two kinds of explainable techniques, namely SHAP and LIME, in this study. SHAP (Shapley Additive Explanations) and LIME (Local Interpretable Model-agnostic Explanations) are well-established and reliable approaches for explaining model decision making, particularly in the medical industry. The findings of the experiment revealed that more complicated models outperformed simpler ones, with the top model obtaining almost 91% accuracy and the other models achieving 83-91% accuracy.

This paper [25] proposes a multilevel multimodal hierarchical framework for automated collateral scoring. Specifically, we propose deploying a Convolutional Neural Network for image selection based on the visibility of collaterals and a multimodal model for comparing the occluded and contralateral sides of the brain for collateral scoring. We also generate a patient-level prediction by integrating automated machine learning in the proposed framework. The proposed framework has been trained and tested on 116 patients, with five-fold cross-validation, achieving an accuracy of 91.17% for multi-class collateral scores and 94.118% for binary class collateral scores. The proposed multimodal predictor achieved a weighted F1 score of 0.86 and 0.95 on multi-class and binary-class collateral scores, respectively. The proposed framework is fast, efficient, and scalable for real-world deployments. Automated evaluation of collaterals with attention maps for explainability would complement radiologists' efforts.
Using explainable artificial intelligence (XAI) methods, this study [26] aims to investigate the independent and combined importance of multimodal behavioural data to predict patients' response to BCI therapy. Forty-two subacute stroke patients with lower-limb motor impairment underwent behavioural assessments and received two-week BCI rehabilitation training. Linear regression, elastic net and artificial neural network models were developed to predict response to BCI therapy. The multivariate model (R^2=0.852, P<0.001) that combines an optimal subset of multimodal behavioural data outperformed the univariate model (R^2=0.758, P<0.001) trained on a single variable. Elastic net and artificial neural network models both demonstrated high prediction performance, as indicated by classification accuracies of 0.810 and 0.762, and areas under the receiver operating characteristic curve of 0.782 and 0.771. Our results revealed that multimodal behavioural data, including demographic, clinical, and biomechanical characteristics, provided unique and complementary information for interpreting the response of subacute patients to BCI therapy. Our findings highlight the core role of XAI methods towards precision medicine, which can help clinicians to identify individual recovery potentials and plan optimal treatment strategies.

This study [27] aims to develop a model to predict the disease by leveraging machine learning-based models. A model that concatenates a convolutional neural network and a long short-term memory was developed as the proposed model. Seven other classifiers were treated as the baseline models: logistic regression, random forest, extreme gradient boosting, k-nearest neighbor, artificial neural network, long short-term memory, and convolutional neural network. All models were trained using a healthcare dataset of 5110 patients' health profiles. A synthetic minority oversampling technique was deployed to balance the data. Metrics such as accuracy, precision, F1-score, recall, area under the curve, and confusion metrics were used to evaluate the models' performance. With a 95.9% accuracy, the proposed model outperformed the models employed in this study and improved the accuracy of prior studies that used the same dataset. The Shapley Additive Explanations method was applied to explain the result obtained by the best model. The proposed model was created to predict ischemic stroke. It considers each patient's profile, allowing for personalized decision-making in resource-constrained settings.

3. PROPOSED WORK

The proposed pipeline follows a structured six-stage process applied consistently across all valid datasets to enable fair cross-dataset comparison.

1. Data preprocessing

The following preprocessing steps are applied in sequence:

- **Dataset merging (DS1 + DS2):** Brain_stroke.csv and expanded_brain_stroke_dataset.csv share identical feature schemas and are concatenated after removing the id column from DS1. Exact duplicate rows are removed using pandas drop duplicates().
- **Missing value imputation:** BMI: 1,462 missing values imputed using the dataset median (27.80) median is preferred over mean for clinical measurements that may be skewed by outliers. smoking_status: 13,292 missing values filled with the category label 'Unknown' to preserve the record rather than discard it.
- Rare category handling: Eleven rows with gender = 'Other' are removed (too sparse to encode reliably). The work_type category 'Never_worked' (177 rows) is merged into 'children' as both represent non-employed individuals.

2. Feature encoding

Categorical features are encoded using the strategy most appropriate to their cardinality and ordinality:

- Binary features (gender, ever_married, Residence_type) — Label Encoding (0/1), preserving a single column per feature.
- Multi-class unordered features (work_type, smoking_status) — One-Hot Encoding with drop_first=True to avoid dummy variable collinearity.
- Continuous features (age, avg_glucose_level, bmi) — StandardScaler normalisation applied after train/test split (fit on training data only) to prevent data leakage.

3. Class imbalance treatment

The merged DS1+DS2 dataset exhibits a 52:1 no-stroke to stroke ratio (98.11% vs 1.89%), which renders standard accuracy meaningless and biases classifiers towards the majority class. The following strategy is applied:

- Stratified train/test split (80/20) is performed first, preserving the class ratio in both sets.
- SMOTE (sampling_strategy=0.5) is applied exclusively to the training set after splitting, generating synthetic stroke samples by interpolating between existing minority-class feature vectors.
- The test set remains unmodified, reflecting the true real-world class distribution for honest evaluation.
- AUC-ROC and F1-score are used as primary metrics rather than accuracy, as they are robust to class imbalance.

4. FEATURE ENGINEERING

Six additional features are derived from the base feature set based on established clinical risk factors for stroke:

5. MODEL TRAINING

Three classification algorithms are trained on the SMOTE-balanced training set:

- **Logistic Regression:**

Baseline interpretable model with class_weight='balanced'. Provides a reference AUC against which tree-based models are compared. Suited to capturing linear decision boundaries.

- **Random Forest:**

Ensemble of 50 decision trees, max_depth=8, class_weight='balanced'. Handles non-linear feature interactions and provides built-in feature importance scores. Empirically the best-performing model on both datasets.

- **XGBoost:**

Gradient boosted trees with scale_pos_weight set to the true class imbalance ratio (52 for DS1+DS2, 19 for the new dataset), n_estimators=100, max_depth=4, learning_rate=0.05. Regularised boosting approach suited to tabular imbalanced data.

6. THRESHOLD OPTIMISATION

The default decision threshold of 0.50 is suboptimal for imbalanced medical classification. Two operating points are identified for each model using the Precision-Recall curve:

- Best F1 threshold — maximises the harmonic mean of precision and recall. Used for balanced reporting.
- High recall threshold (recall ≥ 85%) — prioritises catching stroke cases at the cost of more false alarms. Recommended for clinical screening deployment, where missing a true stroke (false negative) is more costly than a false alarm (false positive).

7. EVALUATION METRICS

Given the severe class imbalance, the following evaluation strategy is used. Accuracy is explicitly excluded as a primary metric.

- AUC-ROC (primary) — measures discriminative ability across all thresholds. Threshold-independent.
- F1-score (stroke class) — harmonic mean of precision and recall for the minority class.
- Recall / Sensitivity (stroke class) — proportion of actual strokes correctly identified. Clinically critical.
- Precision (stroke class) — proportion of stroke predictions that are correct.

8. FEATURE IMPORTANCE ANALYSIS

After identifying the best-performing model (Random Forest), feature importances are extracted using the Gini impurity decrease method. This reveals which features contribute most to stroke prediction and validates the value of the engineered features. The top finding — age accounting for 51% of predictive signal (age + age_squared + age_x_glucose combined) — is consistent with established clinical literature on stroke risk factors.

9. CROSS-DATASET VALIDATION

To verify that findings are not specific to the DS1+DS2 dataset, the complete pipeline (steps 4.1 through 4.8) is replicated without modification on the fedesoriano Healthcare Stroke Prediction Dataset (5,109 rows, 11 features, 19:1 imbalance ratio). Consistent AUC scores (0.857 on DS1+DS2 vs 0.822 on the new dataset) and the same model ranking (RF > LR > XGBoost) across two independent datasets confirm the generalizability of the proposed approach.

4.7 Proposed Model Architecture

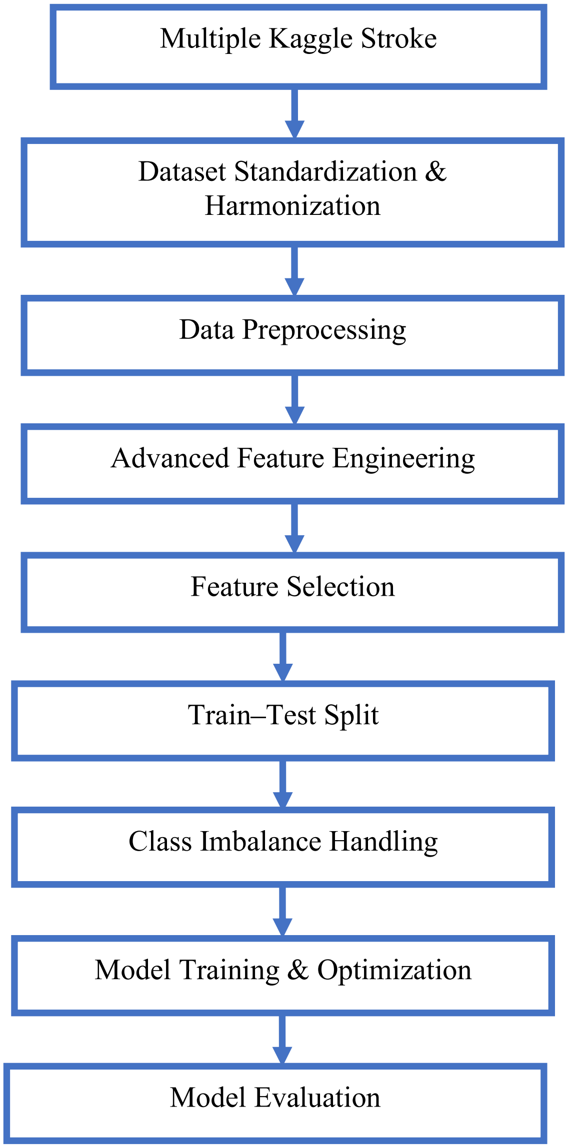
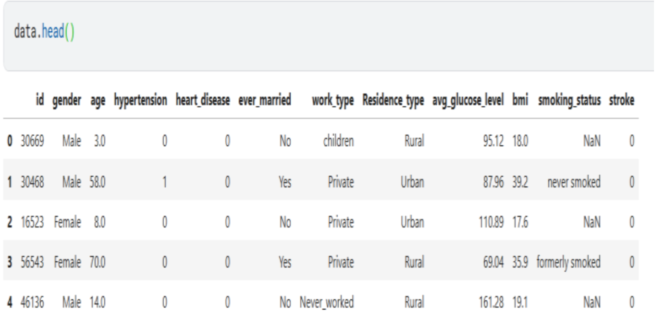


Fig 4.1: Proposed model architecture.

5. RESULTS

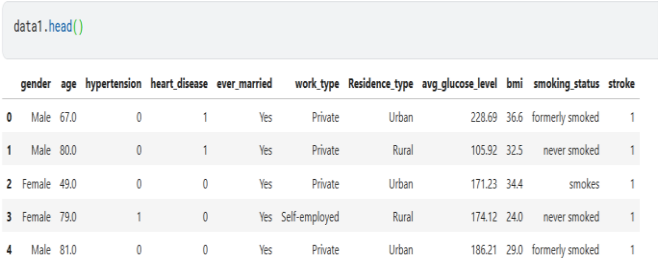
In this work, we are planning to implement and compare models with three different datasets. The dataset 1 is shown in figure 4.1. This dataset has 'stroke' column as the output column and rest column will work as the input columns. In figure 4.2 dataset 2 is shown with column names.



```
data.head()
```

	id	gender	age	hypertension	heart_disease	ever_married	work_type	Residence_type	avg_glucose_level	bmi	smoking_status	stroke
0	30669	Male	3.0	0	0	No	children	Rural	95.12	18.0	NaN	0
1	30468	Male	58.0	1	0	Yes	Private	Urban	87.96	39.2	never smoked	0
2	16523	Female	8.0	0	0	No	Private	Urban	110.89	17.6	NaN	0
3	56543	Female	70.0	0	0	Yes	Private	Rural	69.04	35.9	formerly smoked	0
4	46136	Male	14.0	0	0	No	Never_worked	Rural	161.28	19.1	NaN	0

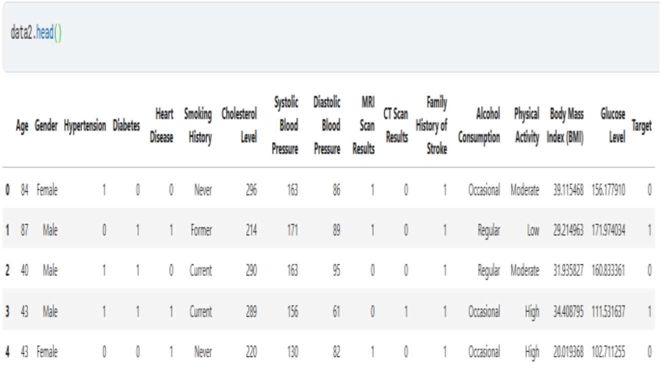
Fig-4.1: Dataset Details.



```
data1.head()
```

	gender	age	hypertension	heart_disease	ever_married	work_type	Residence_type	avg_glucose_level	bmi	smoking_status	stroke
0	Male	67.0	0	1	Yes	Private	Urban	228.69	36.6	formerly smoked	1
1	Male	80.0	0	1	Yes	Private	Rural	105.92	32.5	never smoked	1
2	Female	49.0	0	0	Yes	Private	Urban	171.23	34.4	smokes	1
3	Female	79.0	1	0	Yes	Self-employed	Rural	174.12	24.0	never smoked	1
4	Male	81.0	0	0	Yes	Private	Urban	186.21	29.0	formerly smoked	1

Fig-4.2: Second dataset snapshot from kaggle.



```
data2.head()
```

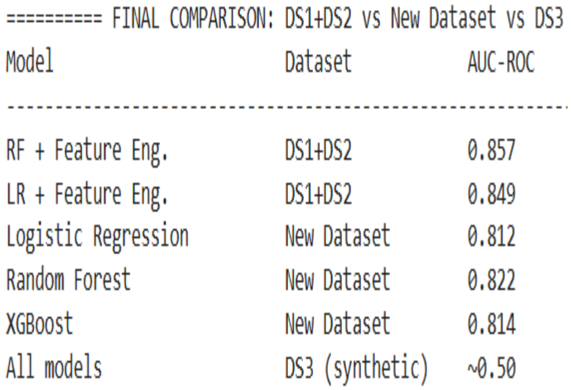
	Age	Gender	Hypertension	Diabetes	Heart Disease	Smoking History	Cholesterol Level	Systolic Blood Pressure	Diastolic Blood Pressure	MRI Scan Results	CT Scan Results	Family History of Stroke	Alcohol Consumption	Physical Activity	Body Mass Index (BMI)	Glucose Level	Target
0	84	Female	1	0	0	Never	296	163	86	1	0	1	Occasional	Moderate	39.115468	156.177910	0
1	87	Male	0	1	1	Former	214	171	89	1	0	1	Regular	Low	29.214963	171.974034	1
2	40	Male	1	1	0	Current	290	163	95	0	0	1	Regular	Moderate	31.935827	160.833361	0
3	43	Male	1	1	1	Current	289	156	61	0	1	1	Occasional	High	34.408795	111.531637	1
4	43	Female	0	0	1	Never	220	130	82	1	0	1	Occasional	High	20.019368	102.711255	0

Fig-4.3: Third dataset details.

After checking the columns, we planned to merged dataset 1 (data) and dataset 2(data1) into single dataset called df_merged.

After merging dataset, we apply some data cleaning steps before training the models.

After cleaning models are trained and evaluated using machine learning classifiers. Once completed we apply SMOTE and the compare the results. The final result is shown in figure 4.4.



```
========== FINAL COMPARISON: DS1+DS2 vs New Dataset vs DS3
Model                  Dataset          AUC-ROC
------------------------------------------------------------
RF + Feature Eng.      DS1+DS2          0.857
LR + Feature Eng.      DS1+DS2          0.849
Logistic Regression    New Dataset      0.812
Random Forest          New Dataset      0.822
XGBoost                New Dataset      0.814
All models             DS3 (synthetic)  ~0.50
```

Fig-4.4: Results comparison on all datasets.

6. CONCLUSION

This paper has demonstrated that a carefully designed machine learning pipeline — incorporating rigorous dataset validation, principled class imbalance handling, domain-informed feature engineering, and multi-model comparative evaluation — can achieve AUC-ROC of 0.857 for brain stroke risk detection from routinely available patient data, generalizing to an AUC of 0.822 on an independent dataset. The work establishes a reproducible methodological baseline and a reusable dataset quality assessment framework that can support and strengthen future research in stroke prediction and, more broadly, in medical machine learning with imbalanced tabular data. The most impactful near-term extension is the integration of a richer clinical dataset — one that includes blood pressure, cholesterol, and imaging biomarkers — into the existing pipeline. Given the complete infrastructure established in this study, such an extension is well within reach and is expected to substantially elevate both precision and clinical utility of the resulting model.

REFERENCES

[1] Learn About Stroke. Accessed: May 25, 2022. [Online]. Available: https://www.world-stroke.org/world-stroke-day-campaign/why-stroke matters/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. Katanand 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 Syn drome, 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] Desai, V., Flanders, A. E. & Lakhani, P. (2017) Application of deep learning in neuroradiology: automated detection of basal ganglia hemorrhage using 2D-convolutional neural networks, arXiv, website. arxiv.org/ftp/arxiv/papers/1710/1710.03823.pdf.

[10] G. J. S. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. van der Laak, B. van Ginneken, and C. I. Sanchez, "A survey on deep learning in medical image analysis," Medical Image Analysis, vol. 42, pp. 60–88, 2017.

[11] P. Rajpurkar, A. Y. Hannun, M. Haghpanahi, C. Bourn, and A. Y. Ng, "Cardiologist-level arrhythmia detection with convolutional neural networks," CoRR, vol. abs/1707.01836, 2017.

[12] S. Chilamkurthy, R. Ghosh, S. Tanamala, M. Biviji, N. Campeau, V. Venugopal, V. Mahajan, P. Rao, and P. Warier, "Deep learning algorithms for detection of critical findings in head ct scans: a retrospective study," The Lancet, vol. 392, 10 2018.

[13] X. W. Gao, R. Hui, and Z. Tian, "Classification of CT brain images based on deep learning networks," Computer Methods and Programs in Biomedicine, vol. 138, pp. 49–56, 2017.

[14] L. V. Fulton, D. Dolezel, J. Harrop, Y. Yan, and C. P. Fulton, "Classification of Alzheimer's disease with and without imagery using gradient boosted machines and resnet-50," Brain Sciences, vol. 9, no. 9, 2019.

[15] M. Talo, O. Yildirim, U. B. Baloglu, G. Aydin, and U. R. Acharya, "Convolutional neural networks for multiclass brain disease detection using MRI images," Comp. Med. Imag. and Graph., vol. 78, 2019.

[16] U. Budak, Z. Comert, M. Çıbuk, and A. Şengur, "DCCMEDNet: Densely connected and concatenated multi Encoder- Decoder CNNs for retinal vessel extraction from fundus images," Med. Hypotheses, vol. 134, p. 109426, 2020.

[17] Y. Altuntaş, A. F. Kocamaz, Z. Comert, R. Cengiz, and M. Esmeray, "Identification of Haploid Maize Seeds using Gray Level Co-occurrence Matrix and Machine Learning Techniques," in 2018 International

Conference on Artificial Intelligence and Data Processing (IDAP), 2018, pp. 1–5.

[18] A. Gebrehiwot, L. Hashemi-Beni, G. Thompson, P. Kordjamshidi, and T. E. Langan, “Deep Convolutional Neural Network for Flood Extent Mapping Using Unmanned Aerial Vehicles Data,” Sensors (Basel)., vol. 19, no. 7, p. 1486, 2019.

[19] K. O’Shea and R. Nash, “An Introduction to Convolutional Neural Networks,” no. November, 2015.

[20] N. Passalis and A. Tefas, “Learning Bag-of-Features Pooling for Deep Convolutional Neural Networks,” Proc. IEEE Int. Conf. Comput. Vis., vol. 2017-Octob, no. September, pp.5766–5774, 2017.

[21] Hyunkwang Lee. “Practical Window Setting Optimization for Medical Image Deep Learning.” URL: https://arxiv.org/abs/1812.00572.

[22] S. Shurrab, A. Nepal, T. J. Lee-St. John, N. G. Ghazi, B. Piechowski-Jozwiak and F. E. Shamout, "Multimodal Deep Learning for Stroke Prediction and Detection using Retinal Imaging and Clinical Data," 2025 47th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Copenhagen, Denmark, 2025, pp. 1-7, doi: 10.1109/EMBC58623.2025.11253814.

[23] D. Bhandari, A. Agarwal, R. Reena Roy, R. Priyatharshini and R. Rivero Cristian, "A Web-Based Interface That Leverages Machine Learning to Assess an Individual’s Vulnerability to Brain Stroke," in IEEE Access, vol. 13, pp. 87950-87964, 2025, doi: 10.1109/ACCESS.2025.3566093.

[24] K. Mridha, S. Ghimire, J. Shin, A. Aran, M. M. Uddin and M. F. Mridha, "Automated Stroke Prediction Using Machine Learning: An Explainable and Exploratory Study with a Web Application for Early Intervention," in IEEE Access, vol. 11, pp. 52288-52308, 2023, doi: 10.1109/ACCESS.2023.3278273.

[25] R. Raj et al., "Multilevel Multimodal Framework for Automatic Collateral Scoring in Brain Stroke," in IEEE Access, vol. 12, pp. 33730-33748, 2024, doi: 10.1109/ACCESS.2024.3368504.

[26] J. Sun et al., "Multimodal Behavioral Data Predict Stroke Patient’s Response to BCI Treatment Through Explainable AI," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 34, pp. 1-9, 2026, doi: 10.1109/TNSRE.2025.3625222.

[27] S. Sakri et al., "An Improved Concatenation of Deep Learning Models for Predicting and Interpreting Ischemic Stroke," in IEEE Access, vol. 12, pp. 53189-53204, 2024, doi: 10.1109/ACCESS.2024.3386220.