# A Hybrid Machine Learning Model in Diagnosing Brain Strokes

Mohammed I. B. Ahmed [1], Rim Zaghdoud [1], Atta Rahman [2,*], Farhan Ali [3,*], Hussain Alhashim [1], Mohammed Y. Almubarak [1], Mohammed Albasheer [1], Abdulwahab Alaqel [1], Ahmed Almaskeen [1], Dina A. Alabbad [1], Danah Aljaafari [4], and Aishah Albakr [4]

[1] Department of Computer Engineering, College of Computer Science and Information Technology, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia
[2] Department of Computer Science, College of Computer Science and Information Technology, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia
[3] College of Electronics and Information Engineering, Shenzhen University, Shenzhen, China
[4] Department of Neurology, College of Medicine, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia
Email: mibahmed@iau.edu.sa (M.I.B.A.); razaghdoud@iau.edu.sa (R.Z.); aaurrahman@iau.edu.sa (A.R.); farhanali@yeah.net (F.A.); 2190002692@iau.edu.sa (H.A.); 2180003080@iau.edu.sa (M.Y.A.); 2190004161@iau.edu.sa (M.A.); 2190004740@iau.edu.sa (A.A.); 2190002369@iau.edu.sa (A.A.); daalabbad@iau.edu.sa (D.A.A.); dtaljaafari@iau.edu.sa (D.A.); abakr@iau.edu.sa (A.A.)
*Corresponding author

***Abstract*—Strokes can occur suddenly and unexpectedly, especially brain strokes, which can be fatal for individuals over the age of fifty. Survivors of a stroke may experience severe paralysis or weakness, posing a significant challenge for healthcare professionals to treat. However, artificial intelligence and Machine Learning (ML) have been proven promising in addressing these critical issues. Despite the high incidence of strokes in countries like Qatar, there is limited research on stroke risk in the Middle East. This study is the first to use a dataset that combines multiple open-source datasets from the region. In this research, several machine learning and ensemble learning algorithms, including Decision Trees (DT), Multiple Layer Perceptron (MLP), K-Nearest Neighbors (KNN), Gaussian Naïve Bayes (GNB), Logistic Regression (LR), Support Vector Machine (SVM), ensemble stacking, and Random Forest (RF) classifier have been investigated. All the algorithms were comprehensively analyzed by tuning their respective hyperparameters using the Grid Search approach and extracting the best features from the dataset through statistical analysis. The proposed ensemble stacking model achieved the highest accuracy and an F1-Score of 98% and 98.29%, respectively. The outcome indicates substantial improvement compared to current approaches in literature with similar datasets.**

***Keywords*—brain stroke diagnosis, ensemble method, Support Vector Machine (SVM), Random Forest (RF), Multiple Layer Perceptron (MLP), stacking, voting**

## I. Introduction

Stroke is a lethal disease with determinantal outcomes; that comes in various types, such as Speech, Face, Arm, and Brain Stroke [1]. S peech Stroke is when a person has speech impairments. At the same time, facial strokes affect the patients' ability to move their facial muscles, such as smiling or rolling their eyes [1]. Arm Stroke occurs when one arm becomes weak or numb, making it difficult to raise it. Brain Stroke happens when neurons do not receive the necessary oxygen and nutrients, which can result in cell death [1]. There are two distinct types of Brain Stroke: Ischemic and Hemorrhagic. The former occurs when blood vessels occlude to a part of the brain, while the latter happens when a blood vessel ikn the brain ruptures, causing bleeding in or around the brain [1].

According to a survey conducted by Feigin *et al*. [2] for the World Stroke Organization (WSO), stroke is one of the fastest-growing diseases worldwide. It is the second leading cause of death and the third leading cause of combined death and disability worldwide (as reported by disability-adjusted life-years lost—DALYs), with an estimated total cost of USD 721 billion (around 0.66% of the global GDP). It was further narrated in the report that during the last three decades, the absolute number of worldwide cases increased significantly [2]:

- 70.0% increase in incident strokes.
- 43.0% of deaths from stroke.
- 102.0% prevalent strokes.
- 143.0% DALYs.
- Overall, 86.0% of deaths and 89.0% of DALYs live in lower-income and Lower-Middle-Income Countries (LMIC).

The latest findings are a wake-up call for Middle East and North African (MENA) countries [3]. The statistics on strokes in the region highlight the urgent need for preventive measures. While some nations have lower rates, the average death rate by stroke in MENA is overwhelming. It's time for the authorities to prioritize awareness campaigns, patient education, and public health initiatives to reduce stroke incidents and save lives.

A statistical study and meta-analysis conducted by

Manuscript received May 26, 2025; revised July 11, 2025; accepted September 2, 2025; published November 25, 2025.

Alqahtani *et al*. [4] focused on stroke cases in Saudi Arabia. They collected data from leading publishers, including PubMed, Scopus, and Web of Science, with a focus on recent and recurring incidents of stroke. The study found that the annual incidence of stroke in Saudi Arabia was 0.029% (95% CI: 0.015–0.047), equivalent to 29 strokes per 100,000 people annually (95% CI: 15–47). Muhammad and Aljohani [5] reported that the annual stroke cases in Saudi Arabia are around 43–57% per 100,000 population, with 85% being ischemic type and 15% hemorrhagic type. Hypertension was identified as the major risk factor behind strokes, followed by diabetes, heart disease, and smoking. After-stroke effects included dysphagia, stress ulcers, pneumonia, depression, anxiety, and abnormal troponin levels. Unfortunately, the mortality rate is 27% higher than in other Middle Eastern countries. Moreover, 95% of stroke patients were treated in non-specialized hospitals. Geographically, Jizan province had the highest number of cases at 26%, followed by the eastern province at 21% and the Riyadh region at 14%. More than 60% of cases were reported in these areas compared to the rest of the kingdom. Basri *et al*. [6] mentioned that there is an increase in stroke cases across the kingdom attributed to ageing residents, inequality of care, lack of self-awareness, and inadequate understanding of strokes.

The study aims to use machine learning to predict and prevent potential brain strokes early. It is apparent from previous research that it is among the most fatal diseases all over the world, especially in the Middle East. Nonetheless, relatively few studies have been conducted in Middle Eastern countries, especially Saudi Arabia. Therefore, this study focuses on patients from the Middle East. The study examines various machine learning and ensemble learning techniques that combine multiple algorithms to achieve improved accuracy. Despite extensive Machine Learning (ML) research on stroke diagnosis, studies focusing on Middle Eastern populations remain scarce, and existing models fail to address regional risk factors like diabetes prevalence. Additionally, the study analyzes data and relationships to gain a deeper understanding of brain stroke symptoms, resulting in more accurate and relevant results. The study was a collaboration between computer scientists and medical professionals. The findings have been reviewed and verified by medical experts in the Department of Neurology to ensure the study's viability and practicality.

The rest of the paper is organized as follows: Section II contains a review of the literature; Section III presents the materials and methods used in the study; empirical studies are reported in Section IV; Section V presents the results and discussion, while Section VI concludes the paper.

## II. Review of Literature

In a study aimed at predicting the Ischemic type of brain stroke, Brugnara *et al*. [7] utilized machine learning principles and techniques. The dataset used in this study was collected from a sequential progression of 246 patients with acute ischemic stroke and large vessel occlusion in the anterior circulation who underwent endovascular treatment between April 2014 and January 2018. Moreover, the study evaluated the accuracy of clinical, multimodal imaging, and angiographic variables in predicting the clinical outcome of endovascular treatment for acute Ischemic stroke. Their relative relevance (mRS-90) was predicted by baseline clinical and traditional imaging parameters with an accuracy of 0.711 (95% CI, 0.705–0.717) and an area under the receiver operating characteristic curve of 0.740 (95% CI, 0.733–0.747).

Heo *et al*. [8] conducted and funded by the National Research Fund of Korea aimed to predict strokes using neural language processing. In this regard, conventional machine learning, as well as deep learning models such as a Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), and Multi-Layer Perceptron (MLP), were used to forecast poor outputs using 5-fold cross-validation and grid search methods. The study achieved 80.7% and 79.9% accuracy for the multi-CNN and CNN models, respectively.

Similarly, Kim *et al*. [9] aimed to predict brain Magnetic Resonance Imaging (MRI) reports with ischemic brain stroke type. This study used 60,564 Computed Tomography (CT) and MRI reports from 17,864 patients. The dataset was split into training and testing sets in a 75:25 ratio, respectively. The highest Area Under the Curve (AUC) value of 0.96 was obtained using Global Vectors (GloVe) for word representation word embeddings with Recurrent Neural Network (RNN).

Govindarajan *et al*. [10] implemented stroke classification by predicting with machine learning algorithms. The dataset was collected from Sugam Multispecialty Hospital in India and included 507 samples. The study achieved 95.3% accuracy using an Artificial Neural Network (ANN) algorithm. Bandi *et al*. [11] attempted to use machine learning principles to build prediction models that could classify the severity of brain strokes. Their dataset consisted of 4,799 samples, with 3,123 males and 1,676 females. They applied various machine learning algorithms, but the Random Forest (RF) algorithm performed best, with an accuracy of 94.23%.

Lehmann *et al*. [12] aimed to predict Ischemic Stroke in Brazil using a dataset of 153 samples from the State University of Londrina's University Hospital. The samples were of both genders and all adults. The researchers achieved an accuracy of 85.2% using a Logistic regression machine learning algorithm. A review article discussed new post-stroke rehabilitation advancements that use wearable technology for data collection and machine learning algorithms for exercise evaluation. The highest accuracy in activity recognition was 99.9%, achieved using a CNN algorithm. In movement classification, the highest accuracy of 99.9% was also achieved by Support Vector Machine (SVM) and ANN [13].

In the past, stroke prognosis forecasting was evaluated from a broad perspective, where baseline non-modifiable characteristics, such as age or stroke severity, were considered the most pertinent determinants. However, several predictive models for disability and mortality in stroke patients have been developed, tested, and validated in recent years. The accuracy of most of these models is about 80% [14].

Sailasya and Kumari [15] aimed to analyze stroke prediction performance using ML classification algorithms. They used a dataset of 5,110 rows and 12 columns, splitting the data into 80% training data and 20% testing data. They achieved a higher accuracy of 82% using the Naïve Bayes (NB) classification. Dev *et al.* [16] aims to detect and predict stroke disease using robust learning approaches. The dataset used contained 5,110 rows and 12 columns, and the highest accuracy achieved was 96% using the RF algorithm. Someeh *et al.* [17] aimed to build a model to predict patients with brain stroke using a deep-learning neural network model. The dataset comprised 332 patients with brain stroke, and the optimal model achieved 99.5% accuracy.

The focus was on a specific type of brain stroke prediction, known as ischemic stroke [18]. Multiple machine-learning algorithms were employed, and the results obtained were carefully observed. The dataset used was constructed from clinically acquired DMI data from 741 samples of both male and female subjects. The research achieved 97% accuracy using a multi-class SVM. In recent studies, advanced machine learning approaches have been used to predict stroke risk. One study used a dataset of 3,254 samples and 11 attributes, including age, gender, hypertension, heart disease, ever married, work type, residence type, average glucose levels, Body Mass Index (BMI), smoking status, and stroke. The study achieved 98% accuracy using ensemble machine learning techniques with algorithms such as Naïve Bayes, random forest, and logistic regression [19].

Another study analyzed various distributed machine learning methods for predicting strokes using the Healthcare Dataset Stroke. A large data platform, Apache Spark, was utilized to conduct this work. ML model performance metrics were calculated using Accuracy, Precision, Recall, and F1-measure. According to the findings, the RF classifier had the highest accuracy, at 90% [20]. The method suggested in this research has successfully decreased the false negative rate while maintaining a relatively high overall accuracy [21, 22]. It translates into a successful reduction in the stroke prediction misdiagnosis rate. According to the novel suggested approach, the false positive rate, accuracy, and sensitivity are predicted to be 33.1%, 71.6%, and 67.4%, respectively.

Cui *et al.* [23] proposed a technique to predict stroke disease, which can occasionally be fatal when certain areas of the brain are not supplied with blood. The ROC area score obtained is 0.94. Before fine-tuning, the values for other performance indicators, including Accuracy, Precision, Recall, and F1-Score, were 0.867, 0.8673, 0.866, and 0.8659, respectively. After fine-tuning, the values were improved to 0.9449, 0.9453, 0.9449, and 0.9448, respectively [24]. This study uses adequately trained machine learning algorithms to prototype a text mining and machine learning-based stroke classification system. Machine learning is a significant predictor and detector in surveillance, medicine, and data management. With a classification accuracy of 95% and a standard deviation of 14.69, artificial neural networks trained using a stochastic gradient descent technique outperformed those trained using other algorithms [25].

Mohammed *et al.* [26] investigated eight machine learning algorithms for stroke classification: K-Nearest Neighbors, Naïve Bayes, Logistic Regression, Decision Tree, Random Forest, Multi-Layer Perceptron (MLP)-NN, Deep Learning, and SVM. The study reveals that Random Forest generates the highest accuracy, precision, recall, and F1-Score of 95.97%, 94.39%, 96.12% and 95.39%, respectively.

These studies show the potential of machine learning approaches in predicting and managing stroke risk, a vital component of everyone's health. Medical records and disease information should be stored systematically to facilitate better analysis and diagnosis of future health risks. Health is regarded as a vital component of everyone's life. Therefore, there is a need for a system of storing information about diseases and their connections. Most disease information can be found in patient case summaries, clinic medical records, and other manually maintained records.

## III. Materials and Methods

This section discusses the used machine learning algorithms in this work. The selection of the algorithms was based on their effectiveness in similar problems in the literature [20–27]. The shortlisted algorithms are MLP, Gaussian Naïve Bayes (GNB), Logistic Regression (LR), Decision Tree (DT), RF, K-Nearest Neighbor (KNN) and Voting.

### A. *Hybrid Synthetic Minority over Sampling Technique*

The basic concept is to increase the number of the few samples while decreasing the number of the most abundant samples to reduce sample imbalance. This method, which combines the Synthetic Minority Over-sampling Technique (SMOTE) and under-sampling algorithms, can solve SMOTE's noise issue [27]. According to the experimental findings, SMOTE performs best with ENN and Tome links.

In Fig. 1, we observe under sampling on the left and how it reduces the number of samples to balance the classes. On the right, we see oversampling and multiplying one class to achieve a balanced dataset. The technique was chosen among the others based on its effectiveness for similar problems in the literature [28–30].

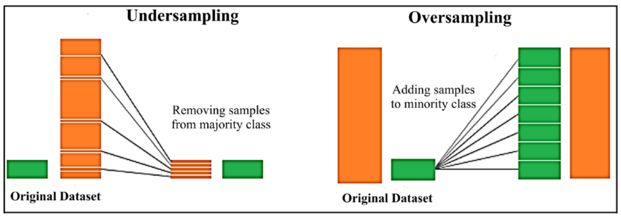


Fig. 1. Oversampling and under-sampling phenomena.

### B. Multi-Layer Perceptron (MLP)

The MLP algorithm is an ANN influenced by the biological neural system of humans [31]. The "neuron", which connects to others through weighted links, is the fundamental component of the MLP. Fig. 2 depicts the elements of an MLP, where we have nodes, input, hidden, and output layers, respectively. It is also regarded as a universal approximator that can approximate any complex phenomena, like [32, 33].

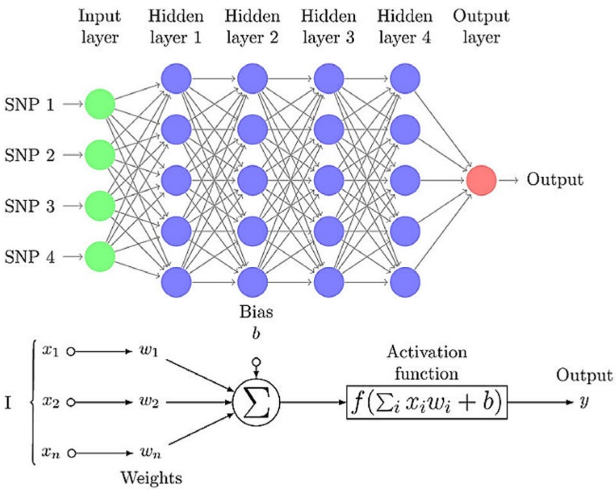


Fig. 2. Multi-Layer Perceptron (MLP) network.

### C. Gaussian Naïve Bayes (GNB)

The GNB is a generalization of the famous NNB algorithm. While many functions can be used to predict data distribution, implementing the Gaussian or normal distribution is a straightforward process because it only requires knowing the mean and standard deviation of the training data [34].

### D. K-Nearest Neighbor (KNN)

KNN is a distance-based machine learning algorithm that works on supervised issues such as classification and regression. An excellent way to understand the algorithm is to visualize a graph with different classification types. For example, the heart disease classification problem is binary and can be classified as positive/negative; and each type will form a block itself [35].

### E. Logistic Regression (LR)

Logistic Regression (LR) is an algorithm for supervised machine learning designed to deal with classification problems. When the target variable is categorical, the issue is called a classification learning problem. It forecasts the probability that a new/unseen example belongs to one of the main target classes. The main objective of the LR algorithm is to map a function from the dataset's features to the targets [36].

### F. Random Forest (RF)

RF models are machine learning techniques that forecast output by combining the results of a series of regression decision trees. Each tree is built separately and is based on a random vector sampled from the input data, with the same distribution of the trees in the forest [37].

### G. Decision Tree

A Decision Tree aims to build a training model that can predict the class or value of a target variable by learning basic decision rules from training data. In Decision Trees, we begin at the tree's root to forecast a class label for a record. We compare the values of the root attribute with the record's attribute. Based on the comparison, it follows the branch corresponding to that value and proceeds to the next node until a leaf node comes, at which point it draws a conclusion [38].

### H. Ensemble Stacking

Ensemble stacking is a technique that combines various classification algorithms to form a superior and more accurate classification model. In stacking, we use multiple classification algorithms to create a meta-model. This higher-level model takes its input from the other prediction models and combines them to produce the best classification result.

It is commonly observed that ensemble stacking has an advantage over other ensemble techniques in yielding higher classification accuracy. However, this comes at the cost of increased computational expense, as the process of combining the predictions from the base classifiers and training the meta-model can be overwhelming [39].

There are no constraints on choosing the input prediction algorithms and the meta-model. Fig. 3 illustrates the process of Ensemble stacking, where Level 0 comprises the base classifiers, which can be simple algorithms, and Level 1 contains the meta-classifier, an advanced algorithm that combines the input obtained from Level 0 [40].

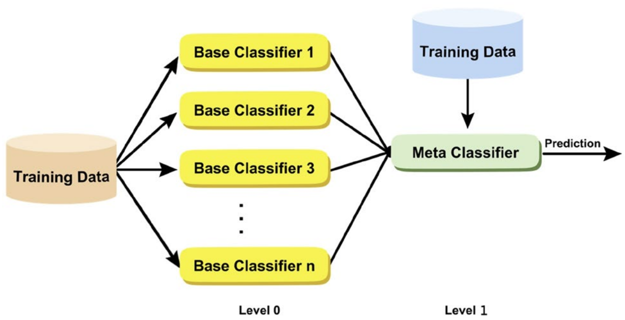


Fig. 3. Ensemble stacking with two levels.

Based on their individual performance, we have shortlisted the models at base and meta models. In the current study, we have employed RF, SVM and LR as based models at Level 0 while MLP is employed as meta model at Level 1. It is based on the literature review feedback as well.

## IV. EMPIRICAL STUDIES

### A. Dataset Description and Statistical Analysis

The dataset was collected from open sources, primarily from Middle Eastern countries as mentioned in its description, including Kaggle [41]. It was also collected from several patients visiting the neurologists there throughout the years, including patients with risk factors such as diabetes and heart disease. The dataset is qualified in terms of demographics, socioeconomic factors, and healthcare systems across the region. Moreover, it was scanned and approved by healthcare professionals from the Department of Neurology at the College of Medicine. The dataset comprises 560 instances, where 410 belong to the negative and 150 belong to the positive class. It is apparent that the dataset is imbalanced, and SMOTE has been employed to balance it. Since only a balanced dataset exhibits a fair analysis. The dataset has already been fused and does not contain the patient's personal information, such as name and contact details. So, no ethical issues are involved. The dataset comprises numerical values with seven attributes, including the target attribute, which is the brain stroke attribute, with two classes: 0, meaning no brain stroke, and 1, meaning a brain stroke. Table I presents the correlation between each attribute and the target attribute. Non numerical attributes are encoded. It reveals the impact of each attribute on the final class outcome. A stronger correlation indicates a higher likelihood of the disease in the presence of the attribute and vice versa. In this regard, diabetes, smoking, heart disease, and hypertension are among the stronger candidate attributes than gender and age, which exhibit a relatively lower correlation coefficient value [41, 42].

TABLE I. CORRELATION WITH TARGET ATTRIBUTE

| Attribute | Correlation coefficient |
|---|---|
| Age | 0.11059407041213658 |
| Gender | 0.003574569466938574 |
| HYPERTENSION | 0.22660857370404536 |
| HEART_DISEASE | 0.235546435885729 |
| SMOKING | 0.3313253719034099 |
| DM | 0.451315437879304 |

TABLE II. STATISTICAL ANALYSIS OF DATASET

| Attribute | Values | Mean | Median | Std. Dev. |
|---|---|---|---|---|
| Age | In years (numeric) | 51.701252 | 51 | 14.884930 |
| Gender | Male (1), Female (0) | Not applicable | | |
| HYPER TENSION | High (1), Low (0) | 0.593918 | - | 0.491540 |
| HEART DISEASE | Positive (1), Negative (0) | 0.098390 | - | 0.298108 |
| SMOKING | Smoker (1), non-smoker (0) | 0.293381 | - | 0.455719 |
| DM (Diabetic) | Positive (1), Negative (0) | 0.422182 | - | 0.494350 |

To conduct the analyses of categorical variables, we performed statistical analysis by SPSS software (version: 26.0, IBM, USA). Quantitative data were compared by independent sample t-test if normal distribution and homogeneity variance requirement were met, or Man-Whitney U test. Categorical data were compared by chi-squared test or Fisher's exact test. The age ($p$ < 0.001 in Man-Whitney U test); gender, Hypertension, Heart disease, Diabetic, and Smoking ($p$ < 0.001 in Chi-squared test) were unbalanced between the two classes (positive vs negative).

The statistical analysis of the dataset is presented in Table II, which includes the data type, mean, median, and standard deviation.

*1) Experimental setup*

Preprocessing transforms the data into a practical and understandable format, making it a fundamental step in classification. It is apparent that quality preprocessing ensures better performance and excellent results. It enables the machine learning algorithm to quickly and accurately extract the unknown features of the data. SMOTE was applied on the training set to balance the dataset.

Subsequently, we divided the dataset into an 80% training set and a 20% test set and employed k-fold cross-validation. Jupyter Notebook and Google Collab were used as the experimental environment, each with built-in machine learning and data analysis tools. The experiments were conducted on a Dell XPS 9320 with 32 GB RAM and a Core i7 architecture.

*2) Performance evaluation*

The performance metric that is most frequently employed in classification tasks will be used in this study. These include the percentage of correctly identified data samples, which are shown by Specificity (SP), Sensitivity (SN), and Accuracy (ACC) are briefly explained below [43–50].

Specificity (*SP*): The percentage of real negative cases that were anticipated to be negative is often referred to as the true negative rate and is determined using the following formula:

$$SP = \frac{TN}{TN + FP} \tag{1}$$

Sensitivity (*SN*): This metric represents the percentage of actual positive cases anticipated to be positive. It is also known as recall. It can be given as:

$$SN = \frac{TP}{TP + FN} \tag{2}$$

Accuracy (*Acc*): This is the proportion of accurate predictions the model made. It comes from:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

F1-Score: It helps to obtain a performance measure by considering all the evaluation factors including True Positive (TP), False Positive (FP), False Negative (FN), and True Negative (TN). F1-Score is measured using the following formula.

$$F1 - Score = \frac{2 \times Recall \times Precision}{Recall + Precision} \tag{4}$$

*3) Hyperparameters tuning*

All the algorithms must be optimized to obtain optimal results in solving the brain stroke classification problem. In this study, we utilized GridSearch to identify the optimal parameters for all the classification algorithms. It is a straightforward approach to addressing the search space of parameters customized by the programmer. It tests all possible parameter combinations and returns the optimal set, which is why GridSearch may take a long time to complete execution when working with substantial-sized data. For instance, for the SVM, the considered hyperparameters in the experiment were cost, kernel, gamma, random state, shrinking, and tol.

Table III lists the optimum values for the hyperparameters obtained and used in the proposed algorithms.

TABLE III. HYPERPARAMETERS VALUES

| Classifier | Hyperparameter | Values Without Sampling |
|---|---|---|
| **SVC** | Cost | 0.1 |
| | Gamma | 0.1 |
| | Kernel | linear |
| | Random state | 0 |
| | Shrinking | True |
| | Tol | 0.001 |
| **KNN** | n_neighbors | 5 |
| | Weights | Distance |
| | Algorithm | Auto |
| | Leaf_size | 20 |
| | P | 1 |
| **MLP** | classifier__hidden_layer_sizes | 100, 100, 100 |
| | classifier_activation | ReLu |
| | classifier_solver | Adam |
| | classifier_alpha | 0.001 |
| **Logistic regression** | C | 0.001 |
| | Max_iter | 500 |
| | Penalty | None |
| | Solver | Sag |
| **Random forest** | Criterion | Gini |
| | Max_depth | 20 |
| | Max_features | None |
| | Min_samples_leaf | 2 |
| | Min_samples_split | 10 |
| | N_estimators | 20 |
| **GNB** | Selector_k | 5 |
| **Decision tree** | Criterion | Gini |
| | max_depth | 10 |
| | min samples split | 20 |
| | min samples leaf | 2 |
| | max features | Auto |
| | Class weight | None |

Regarding feature selection, we extracted and suggested features from the data using guidelines provided by clinical experts, as the data comprised various patients with diverse properties and backgrounds. Based on the clinical data provided, the essential features include smoking, hypertension, heart disease, diabetes, age, and gender, in order from high to low.

### *B. Results and Discussion*

By partitioning the dataset, we have observed a significant effect on accuracy and performance. Partitioning refers to the process of splitting our dataset into training and testing sets. Moreover, 10-fold cross-validation has been performed. Table IV shows the partitioning applied to each split and the associated accuracies. It is apparent from the experiment that the 80–20 split exhibits 98% accuracy, which is 5% more than the 70–30 split, which is 93%. Consequently, we have employed an 80–20 split for the remaining analyses, which utilize all the machine learning algorithms.

TABLE IV. PARTITIONING OF THE DATASET WITH ACCURACIES

| Split ratio | Accuracy of Model stacking |
|---|---|
| 70–30 | 93% |
| 80–20 | 98% |

Table V and Fig. 4 shows the F1-Score values obtained for all proposed techniques.

TABLE V. OBTAINED F1-SCORE FOR THE PROPOSED TECHNIQUES

| Technique | F1-Score |
|---|---|
| Ensemble stacking | 0.9829 |
| Ensemble voting | Hard: 0.72, Soft: 0.79 |
| Random forest | 0.9243 |
| K-Nearest Neighbor | 0.8748 |
| Decision tree | 0.9727 |
| Logistic Regression | 0.9393 |
| Support Vector Machine | 0.9482 |
| Gaussian Naïve Bayes | 0.9649 |
| Multi-layer perceptron | 0.9330 |

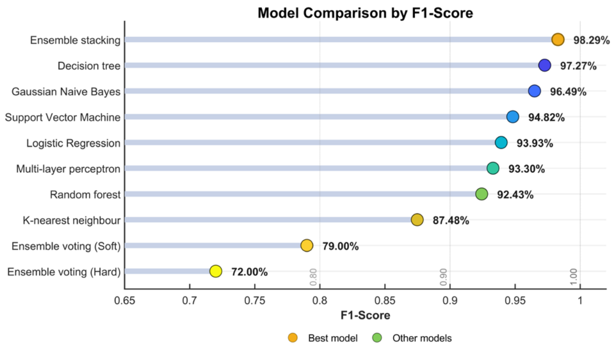


Fig. 4. Obtained F1-Score for the proposed models.

When working with medical problems [51–55], like brain stroke diagnosis, it is crucial to consider the FN score and minimize it as much as possible, which will result in lowering the overall error rate of the models. The results show that Ensemble stacking outperformed all the models, achieving a 98.29% F1-Score. Followed by the Decision Tree at 97.27%, GNB at 96.49%, and SVM at 94.82%, respectively. Likewise, LR and MLP performed almost identically, with rates of 93.93% and 93.30%, respectively. Meanwhile, the KNN performed moderately, achieving an accuracy of 87.48%. Ensemble voting, on the other hand, performed poorly, with 72% (hard) and 79% (soft), respectively. It may be due to the simple nature of the classifier itself, and its constituent voters employed.

## V. FURTHER DISCUSSION

The correlation between the variables along each axis is depicted in each square. The correlation ranges from −1 to +1. Estimates are more akin to zero methods; the two factors exhibit no discernible pattern. Near “1”, the relationship is the more decidedly corresponding; that is, the relationship between the two gets stronger the closer they are to each other. Similar is a correlation closer to −1, but rather than both increasing, one variable will decrease as the other does. Each square is a perfect correlation between each variable and itself, which is why all the diagonals are 1.

Fig. 5 illustrates the heat map generated in the proposed study, highlighting various attributes and their importance and relevance based on their correlation with themselves as well as other attributes. Dark and light colors correspond to strong and poor correlation.


Fig. 5. Heat map for dataset attributes.

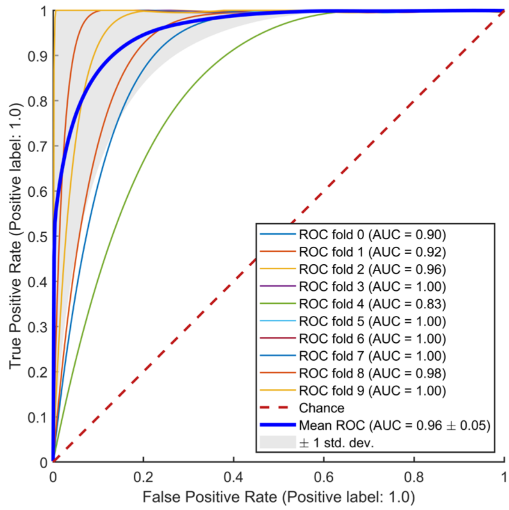


Fig. 6. Receiver operating characteristics for stacking.

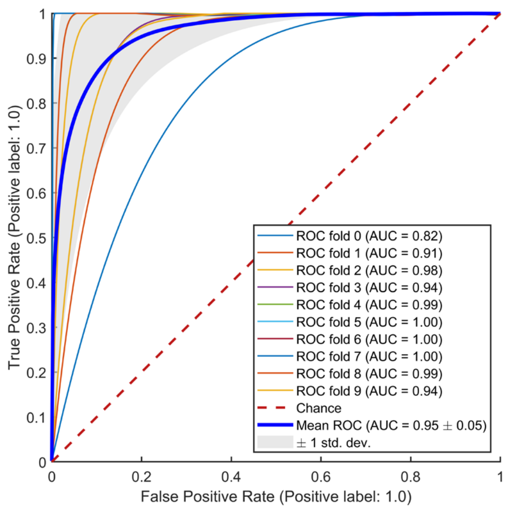


Fig. 7. Receiver operating characteristics for DT.

For further discussion and validation, Receiver Operating Characteristic (ROC) curves are obtained for the top four models, including Stacking, Decision Tree, Gaussian Naïve Bayes, and Support Vector Machine, respectively, in Figs. 6–9. The curves adhere to the performance of each algorithm in terms of its mean ROC values obtained in the given figures. This further warrants the promising nature of the proposed algorithms for brain stroke diagnosis.

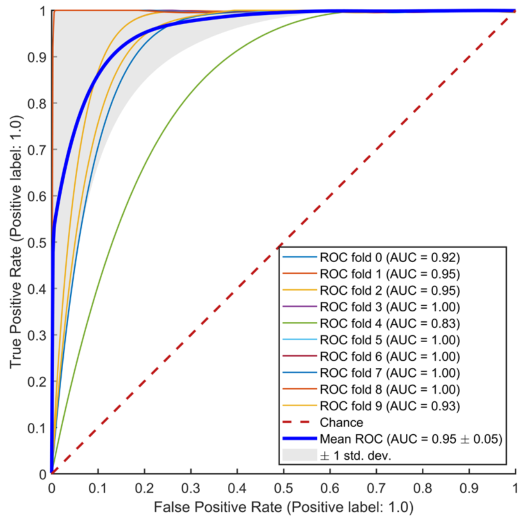


Fig. 8. Receiver operating characteristics for GNB.

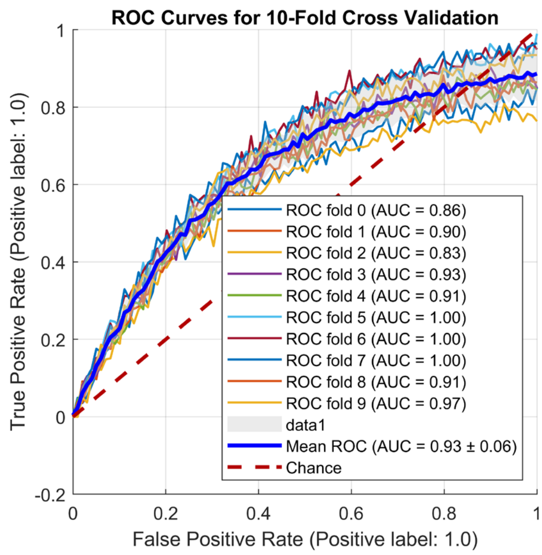


Fig. 9. Receiver operating characteristics for SVM.

Table IV presents the confusion matrix for the best model, which is an ensemble stacking. A false positive rate of 1.43% is observed, which is very acceptable and crucial for medical diagnostic studies using machine learning. Additionally, the precision for the ensemble stacking is observed as 98.21%.

TABLE IV. CONFUSION MATRIX

| | Positive | Negative |
|---|---|---|
| **Positive** | 276 | 4 |
| **Negative** | 7 | 273 |

### A. Weltch's T-Test

In this literature, Welch's t-test is used to test the hypothesis that two populations have equal means and unequal variances, also known as the unequal variance t-test [61]. In the current study, it is reasonable to conduct this test.

Since the test applies to both populations, we combined the diseases as a single population with 279 instances and the no-finding class with 281 cases. Upon calculation [62], the obtained t-value is t = 7.319652. Because the absolute value of the test statistic (4.612) was not larger than the obtained t-value, the null hypothesis of the test cannot be rejected. Hence, there is not sufficient evidence to state that the mean values of the two considered populations are significantly different. The same observation was made by the authors during the data preprocessing phase.

### B. Comparison with State-of-the-Art

This section compares the proposed schemes with state-of-the-art schemes in literature. The schemes were selected based on the region (Middle East), type of stroke, and methods (machine learning) used. They are compared based on two of the most used evaluation metrics: accuracy and F1-Score.

Table VII and Fig. 10 present the comparison with the state-of-the-art. Two schemes were selected on the common grounds, i.e., the same dataset [56, 57]. The comparison was made with the highest accuracy obtained by the scheme using all its best possible approaches. Interestingly, both schemes achieved an accuracy and F1-Score of 96% using the Random Forest algorithm, while the proposed scheme attained an accuracy of 98% and a score of 98.29% using the Ensemble stacking algorithm. The proposed scheme outperformed both by 2% and 2.29% in terms of accuracy and F1-Score, respectively, using the same dataset.

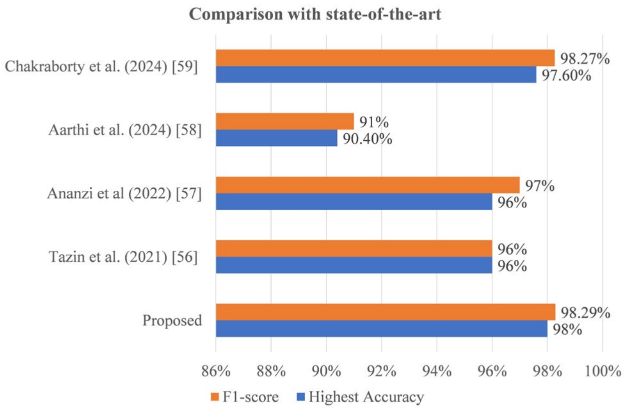


Fig. 10. Comparison with state-of-the-art.

TABLE VII. COMPARISON WITH STATE-OF-THE-ART

| Study | Technique | Accuracy | F1-Score |
|---|---|---|---|
| Proposed | Ensemble stacking | 98% | 98.29% |
| Tazin *et al.* [56] | Random Forest | 96% | 96% |
| Ananzi *et al.* [57] | Random Forest | 96% | 96% |
| Chakraborty *et al.* [59] | Priciple Component Analysis | 97.6% | 98.27% |
| Aarthi *et al.* [58] | Decision Tree | 90.4% | 91% |

The other two recent studies selected for comparison [58, 59], however, utilize an augmented dataset for brain stroke diagnosis. In this regard, the scheme by Chakraborty *et al*. [59] performed more closely to the current study, with marginal differences of 0.4% and 0.02% in accuracy and F1-Score, respectively. However, the scheme by Aarthi *et al*. [58] performed significantly poorly in terms of accuracy and F1-Score, with a difference of 7.6% and 7.29%, respectively. Based on the comparison, it is apparent that the proposed scheme is promising.

### C. Limitations of the Study

The study comprehends the application of various machine-learning approaches on a realistic dataset. The results are promising in accuracy and other evaluation metrics compared to state-of-the-art methods. The scheme is scalable in terms of dataset size and can be adequately extended to larger sizes by utilizing data augmentation techniques.

However, there is always room for improvement. Regarding the study's limitations, the targeted dataset is specific to a particular region. To build a more robust model, data augmentation may be applied, and the results may be re-evaluated. Secondly, the study does not explicitly address the temporal aspect of the data. Changes in lifestyle, healthcare infrastructure, and other external factors over time may influence the accuracy of the diagnosis. Future studies may consider the temporal dynamics of the dataset. Moreover, the study can be extended to incorporate an image dataset to complement the analyses [60].

Explainable Artificial Intelligence (XAI) models, such as Shapley Additive Explanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME), have provided potential insights into the factors influencing disease and diagnosis, along with their ranking [30].

Finally, the study can only predict cases when the patient walks into the hospital and appears to the system for diagnostics. Nonetheless, the study does not provide any mechanism for increasing awareness among the population. Several studies have mentioned that unawareness is among the significant risk factors.

In the future, we are interested in investigating deep learning, transfer learning, and hybrid intelligent models for brain stroke diagnosis on a big dataset by employing data augmentation and preprocessing.

## VI. CONCLUSIONS

This initiative aims to improve healthcare systems using clinical data in the diagnosis of brain strokes and computational intelligence techniques. The early identification of the specified chronic illnesses can be highly beneficial to the health industry and for raising the standard of living in society. It may also slow down their spread, lowering the mortality rate. Therefore, by treating the existing disease, it is advised that this project will positively affect the health and economic sectors. Even though numerous studies were conducted to identify the suggested chronic diseases early on, most of them used imaging technology to gather information that may only be available in some hospitals. Furthermore, no remarkable research mainly targeted datasets from the Middle East, especially Saudi Arabia, to diagnose Brain Stroke in the selected chronic diseases. To create a preemptive diagnosis system that can be used in the hospitals as clinical decision support. By utilizing Middle Eastern clinical datasets that could be used in hospitals with limited resources, this effort is suggested to close the gap to lowering the potential risks associated with the targeted chronic diseases' late discovery. In this regard, a broad spectrum of machine learning algorithms has been investigated in collaboration with healthcare professionals at the Department of Neurology, College of Medicine. Upon conducting a wide range of experiments, it was evident that the proposed schemes could accurately predict the disease up to 98% using the Ensemble staking model. Moreover, the proposed scheme outperforms state-of-the-art techniques regarding accuracy and F1-Score. In the future, we intend to investigate deep learning, ensemble learning, and transfer learning models on diverse datasets, including medical images, such as Magnetic Resonance Imaging (MRI) scans, from Gulf and Middle East regions, to make the model robust and further fine-tuned. Moreover, there is a dire need to build an awareness framework or prototype to bring the population on board to fight this fatal disease.

## Conflict of Interest

The authors declare no conflict of interest.

## Author Contributions

Conceptualization: Mohammed Imran Basheer Ahmed, Danah Aljaafari, and Aishah Albakr. Data curation: Abdulwahab Alaqel. Formal analysis: Mohammed Albasheer. Funding acquisition: Rim Zaghdoud and Atta Rahman. Investigation: Abdulwahab Alaqel, Ahmed Almaskeen and Dina Alabbad. Methodology: Mohammed Imran Basheer Ahmed, Hussain Alhashim, Mohammed Albasheer, Ahmed Almaskeen, Dina Alabbad, Danah Aljaafari and Aishah Ibrahim Albakr. Project administration: Rim Zaghdoud, Farhan Ali, and Atta Rahman. Resources: Atta Rahman and Aishah Albakr. Software: Hussain Alhashim, Mohammed Almubarak, Ahmed Almaskeen and Abdulwahab Alaqel. Supervision: Mohammed Imran Basheer Ahmed, Rim Zaghdoud, and Dina Alabbad. Visualization: Hussain Alhashim, Ahmed Almaskeen, Mohammed Almubarak, and Mohammed Albasheer. Validation: Dina Alabbad and Danah Aljaafari. Revision: Atta Rahman and Farhan Ali. Writing—original draft: Abdulwahab Alaqel, Mohammed Albasheer, and Ahmed Almaskeen. Writing—review & editing: Atta Rahman and Farhan Ali. All authors had approved the final version.

## Acknowledgment

The authors like to acknowledge the collaborative effort among various institutions especially, College of Computer Science and Information Technology (CCSIT) and the College of Medicine at Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia.

## References

[1] Y. Xu, X. Li, D. Wu, Z. Zhang, and A. Jiang, "Machine learning-based model for prediction of hemorrhage transformation in acute ischemic stroke after alteplase," *Frontiers in Neurology*, vol. 13, 897903, 2022.

[2] V. L. Feigin, M. Brainin, B. Norrving *et al*., "World Stroke Organization (WSO): Global stroke fact sheet 2022," *International Journal of Stroke*, vol. 17, no. 1, pp. 18–29, 2022. doi: 10.1177/17474930211065917

[3] M. Jaberinezhad, M. Farhoudi, S. A. Nejadghaderi *et al*., "The burden of stroke and its attributable risk factors in the Middle East and North Africa region, 1990–2019," *Sci. Rep.*, vol. 12, no. 1, 2704, 2022. doi: 10.1038/s41598-022-06418-x

[4] B. A. Alqahtani, A. M. Alenazi, J. C. Hoover *et al*., "Incidence of stroke among Saudi population: A systematic review and meta-analysis," *Neurol. Sci*., vol. 41, no. 11, pp. 3099–3104, 2020. doi: 10.1007/s10072-020-04520-4

[5] A. T. Muhammad and M. M. Aljohani, "Stroke in KSA; epidemiology and clinical delineation," *Majmaah J. Health Sci*., vol. 9, no. 3, pp. 109–120, 2021.

[6] R. Basri, R. Issrani, S. H. Gan, N. Prabhu, and M. K. Alam, "Burden of stroke in the Kingdom of Saudi Arabia: A soaring epidemic," *Saudi Pharm. J.*, vol. 29, no. 3, pp. 264–268, 2021.

[7] G. Brugnara, U. Neuberger, M. A. Mahmutoglu *et al*., "Multimodal predictive modeling of endovascular treatment outcome for acute ischemic stroke using machine-learning," *Stroke*, vol. 51, no. 12, pp. 3541–3551, 2020.

[8] T. S. Heo, Y. S. Kim, J. M. Choi *et al*., "Prediction of stroke outcome using natural language processing-based machine learning of radiology report of brain MRI," *J. Pers. Med*., vol. 10, no. 4, 286, 2020. doi: 10.3390/jpm10040286

[9] C. Kim, V. Zhu, J. Obeid, and L. Lenert, "Natural language processing and machine learning algorithm to identify brain MRI reports with acute ischemic stroke," *PLoS ONE*, vol. 14, no. 2, e0212778, 2019.

[10] P. Govindarajan, R. K. Soundarapandian, A. H. Gandomi *et al*., "Classification of stroke disease using machine learning algorithms," *Neural Comput. & Applic*, vol. 32, no. 1, pp. 817–828, 2020.

[11] V. Bandi, D. Bhattacharyya, and D. Midhunchakkravarthy, "Prediction of brain stroke severity using machine learning," *Revue d'Intelligence Artificielle*, vol. 34, no. 6, pp. 753–761, 2020.

[12] A. L. C. F. Lehmann, D. F. Alfieri, M. C. M. de Araújo *et al*., "Carotid intima-media thickness measurements coupled with stroke severity strongly predict short-term outcome in patients with acute ischemic stroke: a machine learning study," *Metab. Brain Dis*., vol. 36, no. 7, pp. 1747–1761, 2021.

[13] I. Boukhennoufa, X. Zhai, V. Utti, J. Jackson, and K. D. McDonald-Maier, "Wearable sensors and machine learning in post-stroke rehabilitation assessment: A systematic review," *Biomed. Signal Process. Control*, vol. 71, 103197, 2022.

[14] M. Rajora, M. Rathod, and N. S. Naik, "Stroke prediction using machine learning in a distributed environment," in *Proc. ICDCIT*, 2021, pp. 192–203. doi: 10.1007/978-3-030-65621-8_11

[15] G. Sailasya and G. L. A. Kumari, "Analyzing the performance of stroke prediction using ML classification algorithms," *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 6, pp. 539–545, 2021.

[16] S. Dev, H. Wang, C. S. Nwosu *et al*., "A predictive analytics approach for stroke prediction using machine learning and neural networks," *Healthc. Anal.*, vol. 2, 100032, 2022.

[17] N. Someeh, M. A. Jafarabadi, S. M. Shamshirgaran, and F. Farzipoor, "The outcome in patients with brain stroke: A deep learning neural network modeling," *J. Res. Med. Sci.*, vol. 25, 78, 2020. doi: 10.4103/jrms.JRMS_268_20

[18] H. Kamal, V. Lopez, and S. A. Sheth, "Machine learning in acute ischemic stroke neuroimaging," *Front. Neurol*., vol. 9, 945, 2018. doi: 10.3389/fneur.2018.00945

[19] E. Dritsas and M. Trigka, "Stroke risk prediction with machine learning techniques," *Sensors*, vol. 22, no. 13, 4670, 2022. doi: 10.3390/s22134670

[20] H. Ahmed, S. F. Abdelghany, E. M. G. Youn, N. F. Omran, and A. A. Ali, "Stroke prediction using distributed machine learning based on Apache Spark," *Int. J. Adv. Sci. Technol*., vol. 28, no. 15, pp. 89–97, 2019.
[21] A. Rahman, A. K. Luhach, N. Chilamkurti *et al*., "A neuro-fuzzy approach for user behaviour classification and prediction," *J. Cloud Comput*., vol. 8, no. 1, 14, 2019.
[22] T. Liu, W. Fan, and C. Wu, "A hybrid machine learning approach to cerebral stroke prediction based on imbalanced medical dataset," *Artif. Intell. Med*., vol. 101, 101723, 2019. doi: 10.1016/j.artmed.2019.101723
[23] L. Cui, Z. Fan, Y. Yang *et al*., "Deep learning in ischemic stroke imaging analysis: A comprehensive review," *Biomed Res. Int*., vol. 2022, 2456550, 2022.
[24] V. Mariano, J. A. T. Vasquez, M. R. Casu, and F. Vipiana, "Brain stroke classification via machine learning algorithms trained with a linearized scattering operator," *Diagnostics*, vol. 13, no. 23, 3496, 2023.
[25] T. Badriyah, N. Sakinah, I. Syarif, and D. R. Syarif, "Machine learning algorithm for stroke disease classification," in *Proc. ICECCE*, 2020, pp. 1–5.
[26] R. Mohammed, J. Rawashdeh, and M. Abdullah, "Machine learning with oversampling and undersampling techniques: Overview study and experimental results," in *Proc. ICICS*, 2020, pp. 243–248.
[27] T. Sasada, Z. Liu, T. Baba, K. Hatano, and Y. Kimura, "A resampling method for imbalanced datasets considering noise and overlap," *Procedia Comput. Sci.*, vol. 176, pp. 420–429, 2020.
[28] M. Youldash, A. Rahman, M. Alsayed *et al*., "Early detection and classification of diabetic retinopathy: A deep learning approach," *AI*, vol. 5, no. 1, pp. 2586–2617, 2024.
[29] M. N. Alnuaimi, N. S. Alqahtani, M. Gollapalli *et al*., "Transfer learning empowered skin diseases detection in children," *Comput. Model. Eng. Sci.*, vol. 141, no. 3, pp. 2609–2623, 2024. doi: 10.32604/cmes.2024.055303
[30] M. Gollapalli, A. Rahman, S. A. Kudos *et al*., "Appendicitis diagnosis: Ensemble machine learning and explainable artificial intelligence-based comprehensive approach," *Big Data Cogn. Comput*., vol. 8, no. 10, 108, 2024.
[31] G. K. Sharma, S. Kumar, V. Ranga, and M. K. Murmu *et al*., "Artificial intelligence in cerebral stroke images classification and segmentation: A comprehensive study," *Multimed. Tools Appl*., vol. 83, no. 1, pp. 43539–43575, 2024.
[32] A. Rahman, "GRBF-NN based ambient aware realtime adaptive communication in DVB-S2," *J. Ambient Intell. Humaniz. Comput*., vol. 14, pp. 5929–5939, 2023.
[33] A. Rahman, S. Dash, and A. K. Luhach, "Dynamic MODCOD and power allocation in DVB-S2: A hybrid intelligent approach," *Telecommun. Syst*., vol. 76, pp. 49–61, 2021.
[34] R. D. S. Raizada and Y. S. Lee, "Smoothness without smoothing: Why Gaussian naïve Bayes is not naïve for multi-subject searchlight studies," *PLoS ONE*, vol. 8, no. 7, e69566, 2013. doi: 10.1371/journal.pone.0069566
[35] M. I. B. Ahmed, R. A. Zaghdoud, M. Al-Abdulqader *et al*., "Ensemble machine learning based identification of adult epilepsy," *Math. Model. Eng. Probl*., vol. 10, no. 1, pp. 84–92, 2023.
[36] D. Musleh, A. Rahman, M. A. Alkherallah *et al*., "A machine learning approach to cyberbullying detection in Arabic tweets," *Comput., Mater. Contin*., vol. 79, no. 3, pp. 1–22, 2024.
[37] D. A. Musleh, S. O. Olatunji, A. A. Almajed *et al*., "Ensemble learning based sustainable approach to carbonate reservoirs permeability prediction," *Sustainability*, vol. 15, no. 19, 14403, 2023. doi: 10.3390/su151914403
[38] M. I. B. Ahmed, S. Alotaibi, S. Dash *et al*., "A review on machine learning approaches in identification of pediatric epilepsy," *SN Comput. Sci*., vol. 3, no. 6, 437, 2022.
[39] M. G. Meharie, W. J. Mengesha, Z. A. Gariy, and R. N. N. Mutuku, "Application of stacking ensemble machine learning algorithm in predicting the cost of highway construction projects," *Eng., Constr. Archit. Manag*., vol. 29, no. 7, pp. 2836–2853, 2022.
[40] B. Soni. (2023). Stacking to improve model performance: A comprehensive guide on ensemble learning in Python. *Medium*. [Online]. Available: https://medium.com/@brijesh_soni/stacking-to-improve-model-performance-a-comprehensive-guide-on-ensemble-learning-in-python-9ed53c93ce28
[41] Kaggle. (2021). Stroke prediction dataset. [Online]. Available: https://www.kaggle.com/datasets/fedesoriano/stroke-prediction-dataset
[42] S. Arooj, M. F. Khan, T. Shahzad *et al*., "Data fusion architecture empowered with deep learning for breast cancer classification," *Comput., Mater. Contin.*, vol. 77, no. 3, pp. 2813–2831, 2023.
[43] S. M. Alotaibi, A. Rahman, M. I. B. Ahmed, and M. A. Khan, "Ensemble machine learning based identification of pediatric epilepsy," *Comput., Mater. Contin*., vol. 68, no. 1, pp. 149–165, 2021.
[44] A. U. Rahman, A. Alqahtani, N. Aldhafferi *et al*., "Histopathologic oral cancer prediction using oral squamous cell carcinoma biopsy empowered with transfer learning," *Sensors*, vol. 22, no. 10, 3833, 2022. doi: 10.3390/s22103833
[45] D. A. Alabbad, S. Y. Ajibi, R. B. Alotaibi *et al*., "Birthweight range prediction and classification: A machine learning-based sustainable approach," *Mach. Learn. Knowl. Extr.*, vol. 6, no. 2, pp. 770–788, 2024. doi: 10.3390/make6020036
[46] F. Jan, A. Rahman, R. Busaleh *et al*., "Assessing acetabular index angle in infants: A deep learning-based novel approach," *J. Imaging*, vol. 9, no. 10, 242, 2023.
[47] M. S. Ahmed, A. Rahman, F. AlGhamdi *et al*., "Joint diagnosis of pneumonia, COVID-19, and tuberculosis from chest X-ray images: A deep learning approach," *Diagnostics*, vol. 13, no. 15, 2562, 2023. doi: 10.3390/diagnostics13152562
[48] A. Rahman, M. Youldash, G. Alshammari *et al*., "Diabetic retinopathy detection: A hybrid intelligent approach," *Comput., Mater. Contin*., vol. 80, no. 3, pp. 4561–4576, 2024.
[49] W. H. Hantom and A. Rahman, "Arabic spam tweets classification: A comprehensive machine learning approach," *AI*, vol. 5, no. 2, pp. 1049–1065, 2024.
[50] D. A. Alabbad, S. Y. Ajibi, R. B. Alotaibi *et al*., "Birthweight range prediction and classification: A machine learning-based sustainable approach," *Mach. Learn. Knowl. Extr.*, vol. 6, no. 2, pp. 770–788, 2024. doi: 10.3390/make6020036
[51] W. Abbaoui, S. Retal, S. Ziti, B. El Bhiri, and H. Moussif, "Ischemic stroke classification using VGG-16 convolutional neural networks: A study on Moroccan MRI scans," *Int. J. Online Biomed. Eng.*, vol. 20, no. 2, pp. 61–77, 2024.
[52] R. Bakraa, R. Aldhaheri, M. Barashid, *et al*., "Stroke risk factor awareness among populations in Saudi Arabia," *Int. J. Gen. Med*., vol. 14, pp. 4177–4182, 2021. doi: 10.2147/IJGM.S325568
[53] S. Sirisha, S. Jala, S. Vooturi, P. K. Yada, and S. Kaul, "Awareness, recognition, and response to stroke among the general public—An observational study," *J. Neurosci. Rural Pract*., vol. 12, no. 4, pp. 704–710, 2021.
[54] M. O. Owolabi, A. G. Thrift, A. Mahal *et al*., "Primary stroke prevention worldwide: translating evidence into action," *Lancet Public Health*, vol. 7, no. 1, pp. e74–e85, 2022. doi: 10.1016/S2468-2667(21)00230-9
[55] M. M. Qureshi, F. B. Yunus, J. Li, A. Ur-Rahman, T. Mahmood, and Y. A. A. Ali, "Future prospects and challenges of on-demand mobility management solutions," *IEEE Access*, vol. 11, pp. 114864–114879, 2023. doi: 10.1109/ACCESS.2023.3324297
[56] T. Tazin, M. N. Alam, N. N. Dola *et al*., "Stroke disease detection and prediction using robust learning approaches," *J. Healthc. Eng.*, vol. 2021, 7633381, 2021. doi: 10.1155/2021/7633381
[57] E. M. Alanazi, A. Abdou, and J. Luo, "Predicting risk of stroke from lab tests using machine learning algorithms: Development and evaluation of prediction models," *JMIR Form. Res*., vol. 5, no. 12, e23440, 2021. doi: 10.2196/23440
[58] R. Aarthi, P. Vanitha, P. Rajalakshmi, S. J. Thomas, and V. Maadhesh, "Brain stroke prediction using machine learning," in *Intelligent Systems Design and Applications. ISDA 2023*, vol. 1046, 2024, pp. 387–399. doi: 10.1007/978-3-031-64813-7_31
[59] P. Chakraborty, A. Bandyopadhyay, P. P. Sahu *et al*., "Predicting stroke occurrences: A stacked machine learning approach with feature selection and data preprocessing," *BMC Bioinformatics*, vol. 25, no. 1, 329, 2024. doi: 10.1186/s12859-024-05866-8
[60] C. Gao and H. Wang, "Intelligent stroke disease prediction model using deep learning approaches," *Stroke Res. Treat.*, vol. 2024, 4523388, 2024. doi: 10.1155/2024/4523388

[61] H. Nasiri and M. Ebadzadeh, "MFRFNN: Multi-functional recurrent fuzzy neural network for chaotic time series prediction," *Neurocomputing*, vol. 507, pp. 292–310, 2022.