



Publisher homepage: www.universepg.com, ISSN: 2663-7804 (Online) & 2663-7790 (Print)

<https://doi.org/10.34104/ajeit.024.026036>

Australian Journal of Engineering and Innovative Technology

Journal homepage: www.universepg.com/journal/ajeit

Australian Journal of
Engineering and
Innovative Technology



Effective Stroke Prediction using Machine Learning Algorithms

Aklima Begum¹, Md Abdullah Al Mamun^{1*}, Md Atiar Rahman¹, Sabiha Sattar¹, Mist Toma Khatun¹, Shahana Sultana¹, and Mohaimina Begum¹

¹Electronics Division, Atomic Energy Centre, Dhaka-1000, Bangladesh.

*Correspondence: mamun.aec@gmail.com (Dr. Md Abdullah Al Mamun, Principal Scientific Officer, Electronics Division, Atomic Energy Centre, Dhaka-1000, Bangladesh).

ABSTRACT

One of the main factors that lead to death globally is stroke. The main reason for death by stroke is not taking prevention measures early and not understanding stroke. As a result, death by stroke is thriving all over the world, especially in developing countries like Bangladesh. Steps must be taken to identify strokes as early as possible. In this case, machine learning can be a solution. This study aims to find the appropriate algorithms for machine learning to predict stroke early and accurately and identify the main risk factors for stroke. To perform this work, a real dataset was collected from the Kaggle website and split into two parts: train data and test data, and seven machine learning algorithms such as Random Forest, Decision Tree, K-Nearest Neighbor, Adapting Boosting, Gradient Boosting, Logistic Regression, and Support Vector Machine were applied to that train data. Performance evaluation was calculated based on six performance metrics accuracy, precision, recall, F1-score, ROC curve, and precision-recall curve. To figure out the appropriate algorithm for stroke prediction, the performance for each algorithm was compared, and Random Forest was discovered to be the most effective algorithm with 0.99 accuracy, precision, recall, F1-score, an AUC of 0.9925 for the ROC curve, and an AUC of 0.9874 for the precision-recall curve. Finally, feature importance scores for each algorithm were calculated and ranked in descending order to find out the top risk factors for stroke like 'age', 'average glucose level', 'body mass index', 'hypertension', and 'smoking status'. The developed model can be used in different health institutions for stroke prediction with high accuracy.

Keywords: Stroke prediction, Machine learning algorithms, Performance metrics, and Risk factors.

INTRODUCTION:

According to the 'World Health Organization (WHO)', stroke is the 2nd leading cause of mortality around the world and the 1st leading cause in Bangladesh (Global Health Estimates (WHO), 2020). An estimated 11% of all fatalities are caused by strokes, a non-communicable disease (The Top 10 Causes of Death (WHO), 2020). **Fig. 1** shows the latest WHO report on the top 10 leading causes of mortality in Bangladesh over the

year 2019. Therefore, we need to take steps against stroke by making early predictions of its occurrence. Early prediction will help patients take medication before severe conditions occur.

With an increasing number of technologies nowadays, computer-aided systems thrive in the medical sector to diagnose diseases and give treatment (Shehab *et al.*, 2022). Computer-aided systems can work effectively

with high accuracy, and for this reason, they are growing in different sectors (Soofi & Awan, 2017; Hossain *et al.*, 2022). Not only that, but it can also do more work at high speed than any other manual system. Machine learning is a computer-aided system

that is part of artificial intelligence. Machine and learning terms are used to understand machine learning, as machines can learn through a learning process with existing data like humans learn things from experiences (Shinde, 2018).

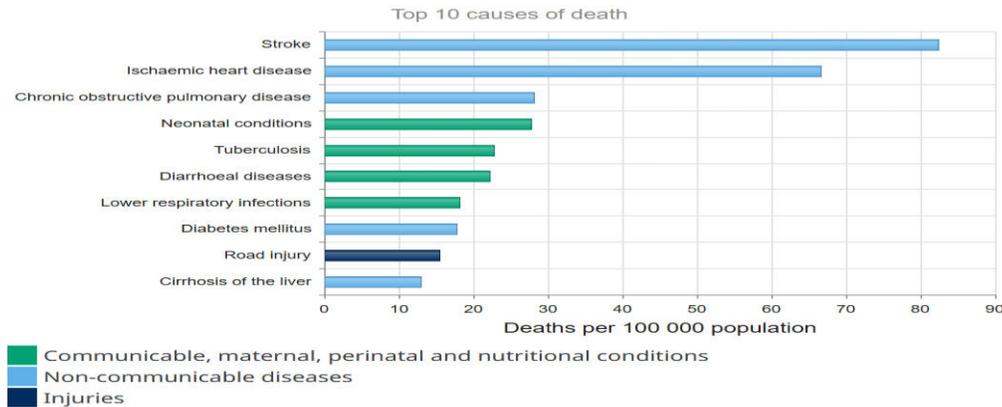


Fig. 1: Leading causes of mortality in Bangladesh (Global Health Estimates (WHO), 2020).

There are several existing algorithms in machine learning that are specially used to develop prediction model (Stiglic *et al.*, 2020; Zhu *et al.*, 2018). Therefore, machine learning algorithms can be a good source to predict strokes as early as possible (Kohli & Arora, 2018). In the medical sector, there are some works based on machine learning algorithms to diagnose diseases like heart disease and cancer (Patel *et al.*, 2016; Seckeler & Hoke, 2011; Shah *et al.*, 2020; Shinde, 2018). However, there are a few studies on stroke prediction with an appropriate algorithm (Emon *et al.*, 2020; Khosla *et al.*, 2010). Another important thing is to find out the main risk factors that trigger the occurrence of stroke, which were less available in previous studies.

In our study, an actual dataset of strokes collected from the Kaggle website to develop an appropriate

model for stroke prediction has been subjected to seven machine learning techniques Random Forest (RF), Decision Tree (DT), K-Nearest Neighbor (K-NN), Adapting Boosting (AdB), Gradient Boosting (XgB), Logistic Regression (LR), and Support Vector Machine (SVM) (Kaggle, 2022). Performance for all applied algorithms was evaluated using different performance metrics and analyzed to find an appropriate algorithm for stroke prediction. After that, feature importance scores were estimated and ranked in descending order to figure out risk factors that trigger stroke.

METHODOLOGY:

To develop a reliable model for stroke prediction, four main steps were counted in this study. The graphical representation of the procedure to develop a prediction model is shown in **Fig. 2**.

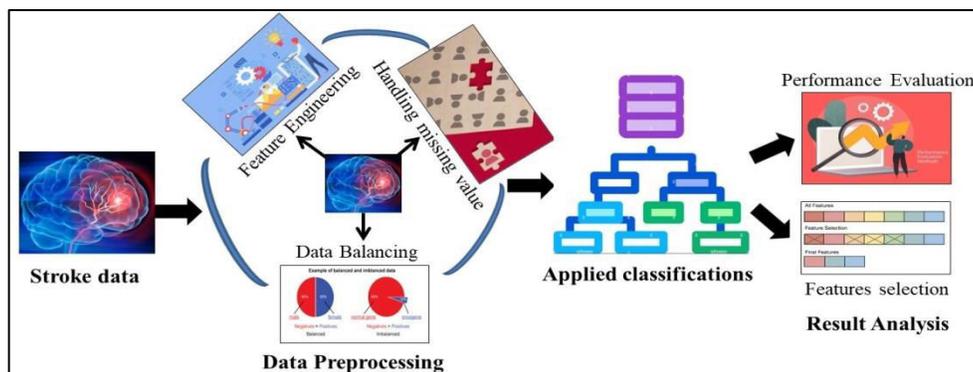


Fig. 2: Graphical visualization of developed model.

Machine Learning Models

This section includes various supervised machine learning algorithms, along with information on how they operate and their benefits, which were applied to develop models to predict stroke.

K-Nearest Neighbor

KNN is a supervised learning which is a slothful algorithm that would not start learning as soon as the dataset was provided. Instead, it holds the dataset and manipulates it when it is time to classify. Finding a resemblance between the latest instance (or data) and the existing data is the foundation of KNN, which then maps the new example into the category that best matches the ones that currently exist (Sailasya & Kumari, 2021).

Random Forest

An RF is made up of numerous distinct decision trees that have each been trained independently using a random subset of data. Each decision tree's output is received during training when these trees are produced. A process known as "voting" is used to determine the algorithm's ultimate prediction. Each decision tree casts a vote for one or more output classes using this procedure. The group with the most votes is selected as the ultimate forecast by the random forest (Sailasya & Kumari, 2021).

Adapting Boosting or AdaBoost

Integrating all of the classes that are weak under an identical powerful classifier is the process known as "boosting." It generates n different multiple decision trees throughout the instruction session. The data record that was first identified erroneously is given top priority after the last decision tree is built. The result associated with these is collected and sent to the subsequent decision model using this. Until we reach the necessary base of learners that was suggested to be created at the beginning, the procedure continues and repeats (Bandi *et al.*, 2020).

Gradient Boosting or XGBoost

This algorithm uses a collection of poor regression models based on decision trees that are constructed in stages, and the resulting model is a composite of those models. The algorithm's ability to ensemble the models and make it possible for efficient use of any

adverse function is its most crucial feature. To put it another way, every regression tree is adapted to the provided loss function's negative gradient, which has been adjusted to the minimal relative variation (Amin Morid *et al.*, 2013).

Decision Tree

The leaflet apex and decision apex are the two apexes of the decision tree, respectively. The Decision apex among them displays the properties of the limbs that are linked to it and are required for decision-making; in difference, these decision limbs' outcomes are represented by the Leaflet apexes, which have no limbs. The choice/ application is contingent on the offered dataset. Graphs are created depending on the required scenarios, showing all the potential outcomes according to the choice or complexity. The tree's structure is estimated as follows since it originates at the source apex and develops in several directions (Bandi *et al.*, 2020).

$$\text{Gini Index (G)} = \sum_{i=1}^c P_i(1 - P_i) \quad (1)$$

Where,

'c' denotes the number of classes, 'pi' denotes the likelihood that class 'i' will exist, and 'G' denotes the root node with the highest value.

Logistic Regression

By using a variable from one of the necessary blocks of independent values, logistic regression is utilized to analyze the absolute dependent values. The LR of the absolute dependent values is usable to analyze the output values. As a result, the solutions can be expressed in terms of absolute or differential variables. It can take on any shape, including binary or numerical variables such as Yes or No, 0 or 1, true or false, etc. The values are presented in a 0 or a format in computer-determined languages, however, this model reflects a viable value that is between 0 and 1. The only distinction between linear regression and logistic regression is how the values are used. While the categorization of the issues was done using LR (Sailasya & Kumari, 2021; Bandi *et al.*, 2020).

$$\text{Sigmoid Function } \phi(z) = \frac{1}{1 + e^{-z}} \quad (2)$$

where,

$z = b_0 + b_1 * \text{age} + b_2 * \text{systolic BP} + b_3 * \text{diastolic BP} + \dots + b_9 * \text{cholesterol}$.

Support Vector Machine (SVM)

The primary goal of SVM is to create a precise linear/deterministic partition that divides n-proportional space into groups to make it simple to combine newly created data into their relevant modules for additional references. Hyperplane is another name for this specific kind of accurate linear data sorting (Sailasya & Kumari, 2021; Bandi et al., 2020). The approach is known as the "Support Vector Algorithm" because it takes into account the outermost vector points, which are support points in the construction of the hyperplane. With the use of demonstration graphs, which are separated into two diverse classes using either a hyperplane or a deterministic division (Bandi et al., 2020).

Proposed Model Architecture

From data collection to final model building, four key processes have been used to develop a prediction model for stroke diagnosis. Firstly, a dataset has been collected from the Kaggle website for stroke with nine features (Kaggle, 2022). Secondly, data has been prepared for the following phase using a variety of Python machine learning modules. Thirdly, to create a model, seven classifiers have been applied and trained on 80% of the data. Lastly, the best stroke prediction model was identified by analyzing the performance of each model. The model architecture is depicted in Fig. 3.

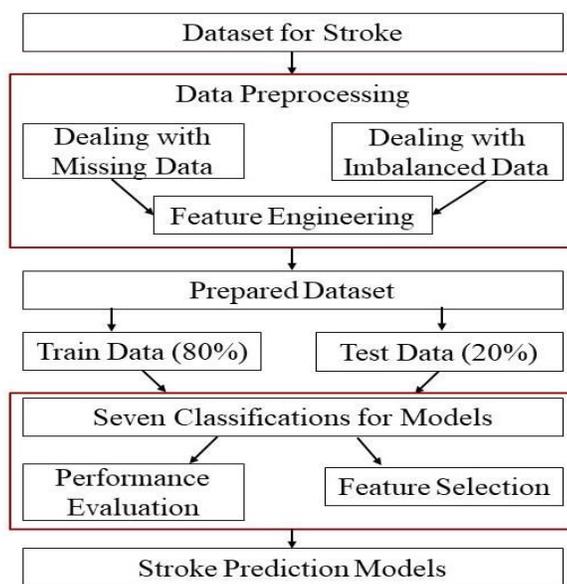


Fig. 3: Model architecture.

Dataset

The dataset for stroke prediction has been collected from Kaggle, which contains ten attributes and a target class of 5110 individuals. Some of these features are

age, ever_married, hypertension, heart_disease, and so on. Table 1 shows the features and the detailed meaning of every feature.

Table 1: Features of stroke dataset with details.

Name of Feature	Meaning	Details
'gender'	Gender	Female/Male, Encoded Female as 1 and Male as 0
'age'	Age	0-deaths, Number in decimal
'hypertension'	Hypertension or Blood Pressure	No hypertension < 140/90 mmHg blood pressure Hypertension = 140/90 mmHg or higher blood pressure (WHO, 2023)
'heart_disease'	Cardiovascular disease	0 = no cardiovascular disease 1 = cardiovascular disease present

'ever_married'	Marital Status	Married/Unmarried Encoded Married as 1 and Unmarried as 0
'work_type'	Type of work	Private/Self-employed/Govt_job Encoded as Private = 0, Self-employed =1, Governmental job (Govt_job) = 2
'residence_type'	Type of residential areas	Urban/Rural Encoded as Urban = 0, Rural = 1
'avg_glucose_level'	The average glucose level in blood	Average blood glucose level in mg/dL
'bmi'	Body Mass Index	BMI < 18.5, underweight BMI = 18.5 to <25, healthy weight BMI = 25.0 to <30, overweight BMI = 30.0 or higher, obesity (CDC, 2022)
'smoking_status'	Status of smoking	Smoker/Nonsmoker Encoded as Smoker = 0, Nonsmoker = 1
'stroke'	Possibility of Stroke	Target class, possibility of stroke occurrence will be predicted

Data Analysis

Data analysis consists of two important parts: one is data preprocessing and another is model training and testing using prepared data.

Data Preprocessing

Generally, the row dataset contains string, imbalance, and missing data which are not be proceed in machine learning algorithms. In this study, 'gender', 'residence_type', 'work_type', and 'ever_married' were in string mode in the dataset as shown in **Table 1**, which were needed to be converted into numbers, and Label Encoder has been used to perform this task. In the given dataset, there were 201 nan values, or missing values, which means 201 values of different features were invalid or empty. Machine learning algorithms cannot perform on these nan values. So, the SMOTE technique has been used to fill up the missing values with the mean or median value of other values of the same feature. The row stroke dataset contained 249 target classes 0 and 4861 target classes 1, which means the dataset was imbalanced. With this imbalance, the model cannot provide actual results. So, the over sampling technique has been utilized to overcome this problem, and after oversampling both class data, 4861.

Model Training and Testing

After preprocessing, the dataset has become ready for usage. Using the Cross-Validation technique, it was divided into train and test datasets, with train data making up 80% of the total dataset and test data

making up 20%. In this work, models were trained using train data, and model performance was evaluated using testing data.

Performance Evaluation Metrics

The most significant phase in creating a successful model in machine learning is evaluating its performance. To measure the effectiveness of the developed model, different metrics called performance metrics are used (Javatpoint, 2021; Erickson & Kitamura, 2021). After completing classification using train data, results found a groups or classes. Through this training process model gains knowledge to classify new data into a group or class. As the result, it foresees class labels like Yes or No, Spam or Not Spam, 0 or 1, etc. A variety of metrics are used to assess a categorization model's performance (Erickson & Kitamura, 2021; Javatpoint, 2021). The following metrics were used to measure model performance in our study:

- ✓ Accuracy
- ✓ Precision
- ✓ Recall
- ✓ F1-score
- ✓ ROC curve
- ✓ Precision-recall curve

Accuracy

One of the easiest performance metrics to deploy model is accuracy and it is calculated as the ratio of successful predictions to all predictions (Erickson & Kitamura, 2021).

$$\text{Accuracy} = \frac{\text{Number of correct prediction}}{\text{Total number of prediction}} \quad (3)$$

Precision

The precision determines the percentage of positive accurate predictions. It is calculable as the proportion of all true positive (True Positive-TP and False Positive-FP) forecasts to the total positive predictions (Erickson & Kitamura, 2021). Confusion metrics can be used to get true positive and false positive predictions number.

$$\text{Precision} = \frac{TP}{TP+FP} \quad (4)$$

Recall

“It is comparable to the Precision metric and attempts to determine the percentage of actual positives that were mistakenly detected. It can be measured as True Positive, or forecasts that actually match the overall number of positives, either correctly forecasted as positive or wrongly anticipated as negative (true Positive and false negative)”(Javatpoint, 2021). The formula to calculate this is given below (Erickson & Kitamura, 2021).

$$\text{Recall} = \frac{TP}{TP+FN} \quad (5)$$

F1-score

“F1-score, also known as F-Score, is a statistic used to assess a binary classification model based on predictions provided for the positive class. It is a particular kind of score that combines Precision and Recall. As a result, the F1 Score can be determined by taking the harmonic mean of both precision and recall and giving each variable equal weight” (Erickson & Kitamura, 2021). F1-score is the good option to check performance of classifiers rather than accuracy (Orozco-Arias et al., 2020).

$$\text{F1-score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (6)$$

ROC curve

ROC is a abbreviation of “Receiver Operating Characteristic curve” (Mandrekar, 2010). The performance of a classification model across different threshold levels is represented graphically by the ROC curve. True Positive Rate (TPR) and False Positive Rate (FPR) are the two parameters that are utilized to plot this curve. The following equations are used to estimate TPR and FPR (Erickson & Kitamura, 2021).

$$\text{TPR} = \frac{TP}{TP+FN} \quad (5) \quad \text{and} \quad \text{FPR} = \frac{FP}{FP+TN} \quad (7)$$

Area under this curve called AUC and the performance of models are identified based on the value of AUC. AUC determines performance for all thresholds and offers an overall measurement. AUC stands for “area under the curve”(Javatpoint, 2021). Different values of AUC mean different level of model performance are given below:

AUC under 0.7; the developed model is not acceptable;

AUC = 0.7 to 0.8; the developed model is acceptable;

AUC = 0.8 to 0.9; the model performance is excellent;

AUC above 0.9; the model performance is outstanding (Mandrekar, 2010).

Precision-recall curve

Using confusion matrix, precision, specificity and recall can be measured. To produce the Precision-recall (PR) curve every pair of observations at each threshold are connected like ROC curves. PR curve graphs are mainly precision which means truly positive predicted value versus recall (Miao & Zhu, 2022). In case of imbalance data, precision-recall is better performance metric than ROC to estimate good models (Orozco-Arias et al., 2020). Area under this curve (AUC) can be measured to evaluate the performance of models. A PRC that goes through the upper right corner (representing 100% recall and 100% precision and AUC = 1.0) characterizes the ideal classifier. It known that the better the model is, the nearer a point is to that place (Miao & Zhu, 2022).

RESULTS AND DISCUSSION:

To better comprehend how attributes relate to one another, a correlation heatmap was created, with deeper colors denoting weak associations and lighter colors denoting significant correlations. There were also two more possibilities: positive correlation, which occurs when one variable rises as other variables rise, and negative correlation, which occurs when one variable fall while other variables rise. **Fig. 4** shows the correlation among the attributes of stroke dataset. To measure the performance of each model, six performance metrics were used, such as accuracy, precision, recall, F1-score, ROC curve, and precision-recall curve. Table 2 illustrates the basic four performance

criteria for seven supervised learning classifiers (DT, K-NN, RF, LR, SVM, AdB, and XgB). Comparing these classifiers, RF has been acted as the best-performing model with 0.99 accuracy, precision, recall, and F1-score and DT as the second-best model with 0.98 accuracy, precision, recall, and F1-score, while

SVM’s performance was the lowest with 0.77 accuracy, precision, recall, and 0.76 F1-score and LR has also been acted as low-performing algorithms with 0.77 accuracy, precision, recall, and F1-score.



Fig. 4: Correlation heatmap for the features of stroke dataset.

Table 2: Performance metrics for each model.

Performance metrics	DT	K-NN	RF	LR	SVM	AdB	XgB
Accuracy	0.98	0.92	0.99	0.77	0.77	0.81	0.97
Precision	0.98	0.92	0.99	0.77	0.77	0.81	0.97
Recall	0.98	0.90	0.99	0.77	0.77	0.81	0.97
F1-score	0.98	0.90	0.99	0.77	0.76	0.80	0.97

A graphical visualization was also generated to compare the accuracy of each model. Fig. 5 shows the comparison among the applied classifiers to figure out the best-performing model with the appropriate algorithm. According to the result, RF performed well, with the highest accuracy score.

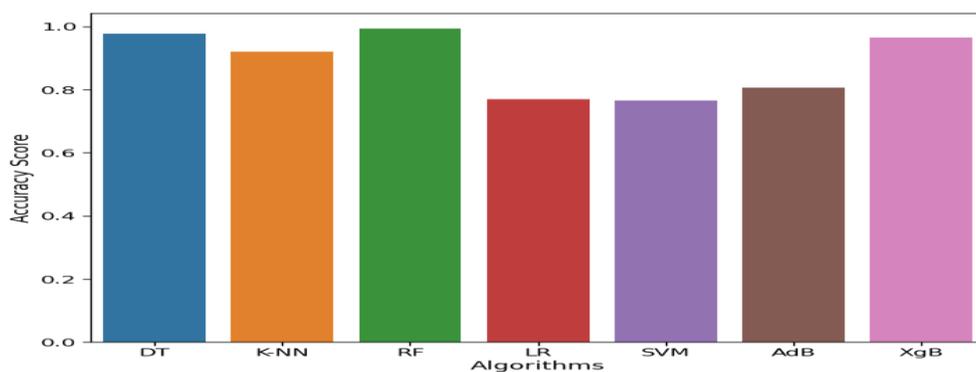


Fig. 5: Comparison among stroke prediction models.

Table 3 compares the recent studies with our study. From Table 3, it is revealed that our study provided

better performance than previous studies, and RF acted as the best-performing algorithm.

Table 3: Comparison among previous studies and our study on stroke prediction.

Previous studies	Models	Accuracy (%)	Precision (%)	Recall (%)
(Bandi et al., 2020)	DT	93.12	91.84	93.75

	LR	82.50	93.23	62.83
	SVM	83.22	95.33	61.06
	RF	96.97	94.56	97.81
(Sailasya & Kumari, 2021)	DT	66	77.5	77.5
	LR	78	77.5	77.6
	SVM	80	78.6	83.8
	RF	73	72	73.5
	K-NN	80	77.4	83.7
Our study	DT	98	98	98
	K-NN	92	92	90
	RF	99	99	99
	LR	77	77	77
	SVM	77	77	77
	AdB	81	81	81
	XgB	97	97	97

Another important performance criterion is the ROC curve. The AUC value defines the performance of a model. The value of AUC under 0.7 means the developed model is not acceptable; from 0.7 to 0.8 means the model performance is good; from 0.8 to 0.9 means excellent; and above 0.9 means outstanding (Mandrekar, 2010). In the medical sector, it is required that the value of the AUC be above 0.9 to prevent

unintentional diagnostic mistakes. Fig. 6 illustrates the ROC curves and their corresponding AUC values for the applied algorithms. Comparing the AUC values of ROC curves, it is shown that RF, DT, XgB, and K-NN give an AUC above 0.9, and RF gives the highest AUC of 0.9925. Therefore, it can be concluded that the RF acts as the best-performing algorithm.

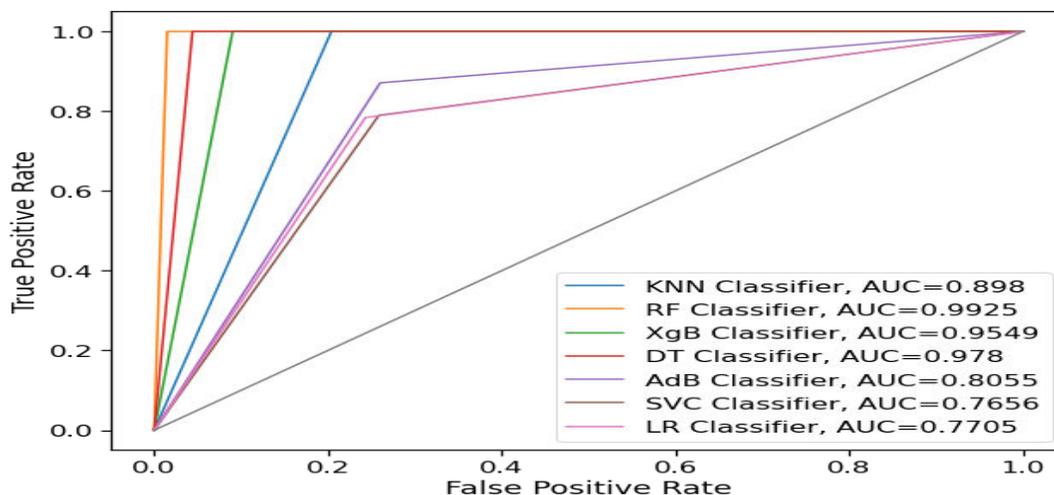


Fig. 6: Comparison among ROC curves for stroke prediction models.

In our study, the PR curve has been used as the final performance metric, which works better than the ROC curve (Orozco-Arias et al., 2020). Fig. 7 shows the PR curves with AUC values. According to the following In Fig. 7, RF gives the highest AUC value of 0.9874 for the PR curve, and DT and XgB also provide AUC values above 0.9. So, it is concluded that RF, DT, and XgB can be used as the appropriate algorithms for stroke prediction models, and among these, RF is the

best-performing algorithm. To find out the main risk factors for stroke, feature importance scores for all applied algorithms were calculated. After finding the feature scores, the features for all algorithms were ranked in descending order of score. Figure 8 shows the graphical visualization of the top five important attributes. Stroke occurring mainly depends on these attributes, which means these are the main risk factors for stroke.

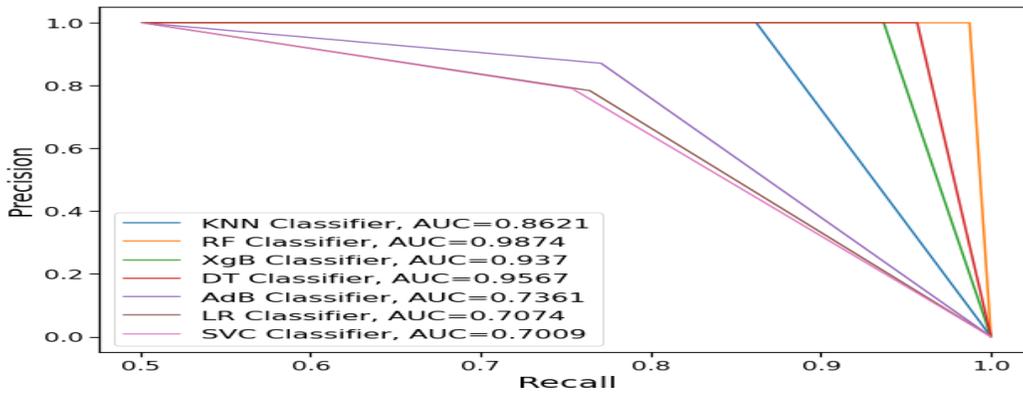


Fig. 7: Comparison among PR curves of stroke prediction models.

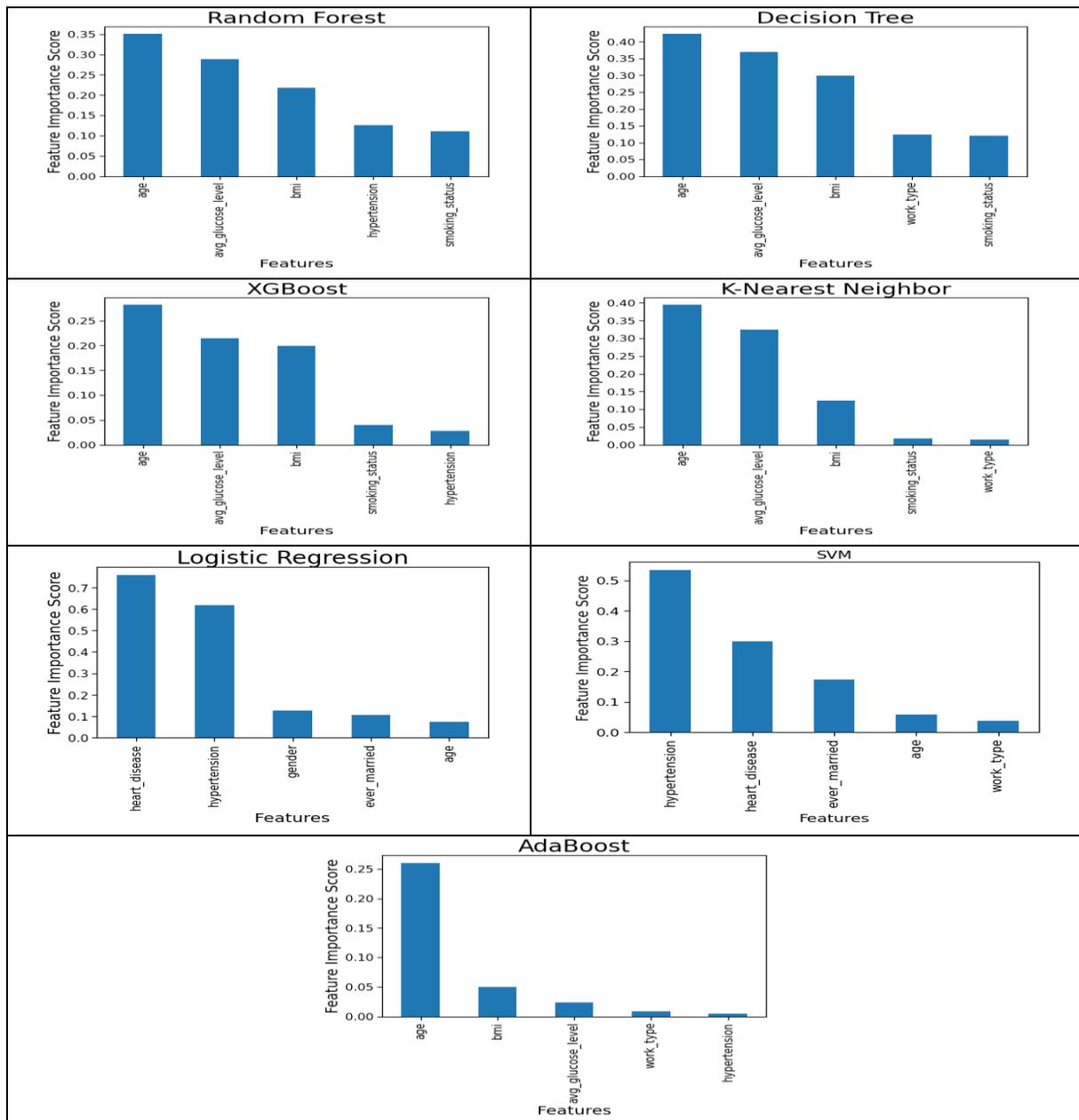


Fig. 8: Top risk factors of stroke for each model

CONCLUSION:

Seven machine learning techniques were applied to a real dataset to create a stroke prediction model, including Random Forest, Decision Tree, K-Nearest Neighbor, Adapting Boosting, Gradient Boosting, Logistic Regression, and Support Vector Machine. The performance of all algorithms employed was then assessed using six performance criteria. After comparing the performance of each algorithm, it was found that Random Forest had the highest accuracy, precision, recall, F1-score, AUC of the ROC curve, and AUC of the precision-recall curve. It was also the best-performing algorithm overall. The scores of feature importance for all algorithms were used to determine the top risk factors. According to the scores, 'age', 'avg_glucose_level', 'bmi', 'hypertension' and 'smoking_status' were the high-risk factors that needed to be treated. The developed model using Random Forest can be used as a decision-making model in the medical sector for the prediction of stroke.

ACKNOWLEDGEMENT:

The authors acknowledge the supports received from the Electronics Division, Atomic Energy Centre, Dhaka-1000, Bangladesh.

CONFLICTS OF INTEREST:

The authors declare that they have no known conflict of interests.

REFERENCES:

- 1) Amin Morid, M., Kawamoto, K., and Abdelrahman, S. (2013). Utah Health Plans for. 1312–1321. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5977561/pdf/2731392.pdf>
- 2) Bandi, V., Bhattacharyya, D., & Midhunchakravathy, D. (2020). Prediction of Brain Stroke Severity Using Machine Learning. *Revue d'Intelligence Artificielle*, 34(6), 753-761. <https://doi.org/10.18280/ria.340609>
- 3) Defining Adult Overweight & Obesity | Overweight & Obesity | CDC. (2022). Retrieved September 6, 2023, from - <https://www.cdc.gov/obesity/basics/adult-defining.html>
- 4) Emon, M. U., Keya, M. S., & Kaiser, M. S. (2020). *Performance Analysis of Machine Learning Approaches in Stroke Prediction*. 1464-1469. <https://doi.org/10.1109/ICECA49313.2020.9297525>
- 5) Erickson, B. J., & Kitamura, F. (2021). Magician's corner: 9. performance metrics for machine learning models. *Radiology: Artificial Intelligence*, 3(3). <https://doi.org/10.1148/RYAI.2021200126>
- 6) Global health estimates: Leading causes of death, (2020). Retrieved September 8, 2023, from - <https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates/ghle-leading-causes-of-death>
- 7) Hossain MA, Hossen E, and Asraful M. (2022). Study and prediction of covid-19 cases and vaccinations using machine learning in Bangladesh. *Aust. J. Eng. Innov. Technol.*, 4(6), 130-139. <https://doi.org/10.34104/ajeit.022.01300139>
- 8) Hypertension. (n.d.). Retrieved September 6, 2023, from <https://www.who.int/news-room/fact-sheets/detail/hypertension>
- 9) Khosla, A., Cao, Y., and Lee, H. (2010). An integrated machine learning approach to stroke prediction. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 183-191. <https://doi.org/10.1145/1835804.1835830>
- 10) Kohli, P. S., & Arora, S. (2018). Application of Machine Learning in Disease Prediction. *Inter Conference on Computing Communication and Automation (ICCCA)*, 1-4. <https://doi.org/10.1109/CCAA.2018.8777449>
- 11) Mandrekar, J. N. (2010). Receiver operating characteristic curve in diagnostic test assessment. *J. of Thoracic Oncology*, 5(9), 1315-1316. <https://doi.org/10.1097/JTO.0b013e3181ec173d>
- 12) Miao, J., and Zhu, W. (2022). Precision–recall curve (PRC) classification trees. *Evolutionary Intelligence*, 15(3), 1545-1569. <https://doi.org/10.1007/s12065-021-00565-2>
- 13) Orozco-Arias, S., Piña, J. S., & Isaza, G. (2020). Measuring performance metrics of machine learning algorithms for detecting and classifying transposable elements. *Processes*, 8(6). <https://doi.org/10.3390/PR8060638>

- 14) Patel, J., Upadhyay, T., & Patel, S. (2016). Heart Disease Prediction Using Machine learning and Data Mining Technique. *International Journal of Computer Science & Communication*, 7(March), 129-137.
<https://doi.org/10.090592/IJCSC.2016.018>
- 15) Performance Metrics in Machine Learning - Javatpoint, (2021). Retrieved September 8, 2023, from
<https://www.javatpoint.com/performance-metrics-in-machine-learning>
- 16) Sailasya, G., & Kumari, G. L. A. (2021). Analyzing the Performance of Stroke Prediction using ML Classification Algorithms. *12*(6), 539-545.
<https://dx.doi.org/10.14569/IJACSA.2021.0120662>
- 17) Seckeler, M. D., & Hoke, T. R. (2011). The worldwide epidemiology of acute rheumatic fever and rheumatic heart disease. *Clinical Epidemiology*, *3*(1), 67.
<https://doi.org/10.2147/CLEP.S12977>
- 18) Shah, D., Patel, S., & Bharti, S. K. (2020). Heart Disease Prediction using Machine Learning Techniques. *SN Computer Science*, *1*(6), 345.
<https://doi.org/10.1007/s42979-020-00365-y>
- 19) Shehab, M., Abualigah, L., & Gandomi, A. H. (2022). Machine learning in medical applications: A review of state-of-the-art methods. *Computers in Biology and Medicine*, *145*, 105-458.
<https://doi.org/10.1016/j.compbiomed.2022.105458>
- 20) Shinde, P. P. (2018). A Review of Machine Learning and Deep Learning Applications. 2018 *Fourth International Conference on Computing Communication Control and Automation (ICC-UBEA)*, 1-6.
<https://doi.org/10.1109/ICCUBEA.2018.8697857>
- 21) Soofi, A. A., & Awan, A. (2017). Classification Techniques in Machine Learning: Applications and Issues. *Journal of Basic & Applied Sciences*, *13*, 459-465.
<https://doi.org/10.6000/1927-5129.2017.13.76>
- 22) Stiglic, G., Kocbek, P., & Cilar, L. (2020). Interpretability of machine learning-based prediction models in healthcare. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *10*(5), 1-13.
<https://doi.org/10.1002/widm.1379>
- 23) Stroke Prediction Dataset | Kaggle. (2022). Retrieved September 6, 2023, from -
<https://www.kaggle.com/datasets/fedesoriano/stroke-prediction-dataset>
- 24) The top 10 causes of death. (2020). Retrieved September 6, 2023, from -
<https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death>
- 25) Zhu, D., Cai, C., & Zhou, X. (2018). A machine learning approach for air quality prediction: Model regularization and optimization. *Big Data and Cognitive Computing*, *2*(1), 1-15.
<https://doi.org/10.3390/bdcc2010005>

Citation: Begum A, Mamun MAA, Rahman MA, Sattar S, Khatun MT, Sultana S, and Begum M. (2024). Effective stroke prediction using machine learning algorithms. *Aust. J. Eng. Innov. Technol.*, *6*(2), 26-36.

<https://doi.org/10.34104/ajeit.024.026036> 