OPEN ACCESS
International Journal of Innovations in Science & Technology

Advancing Diagnosis Capabilities with Smart AI Techniques for Early Symptoms Prediction of Brain Stroke

Muhammad Aurangzeb Khan[1], Raza Ullah Khan[2], Rasool Jamal[2], Zahid Ullah[3], Amjad Khan[1], Farzeed Khan[1]
[1]Department of Computer Science, University of Science and Technology Bannu, KP, Pakistan.
[2]Department of Medicine, Lady Reading Hospital Peshawar, KP, Pakistan.
[3]Department of Cardiology, Lady Reading Hospital Peshawar, KP, Pakistan.
***Correspondence**: farzeedkhanmcs@gmail.com

Citation | Khan. M. A, Khan. R. U, Jamal. R, Ullah. Z, Khan. A, Khan. F, "Advancing Diagnosis Capabilities with Smart AI Techniques for Early Symptoms Prediction of Brain Stroke", IJIST, Vol. 07 Issue. 01 pp 422-439, Feb 2025
Received | Jan 19, 2025 **Revised |** Feb 21, 2025 **Accepted |** Feb 22, 2025 **Published |** Feb 23, 2025.

The brain, a vital organ in the human body, can suffer severe damage during a Brain Stroke (BS) due to blocked blood vessels. The interruption in blood flow and nutrient supply leads to significant symptoms and is considered a medical emergency. BS often results in long-term neurological impairments, complications, or even death, underscoring its critical nature. The World Health Organization (WHO) estimates that BS is the most prevalent cause of disability and death globally. Failure to detect a stroke early may result in delayed treatment, leading to severe complications such as lifelong neurological impairment or death. Early identification with Machine Learning (ML) and Deep Learning (DL) approaches can improve the treatment of patients and reduce the long-term impacts of stroke. The purpose of this research is to predict the signs of a stroke taking place at an early stage employing ML and DL models. To evaluate the efficiency of the approach, a comprehensive training set for BS recognition was collected from a well-known source, Kaggle. The training dataset contains eleven attributes, including age, gender, hypertension, etc., with 5110 records. Multiple classification models, like Support Vector Machine (SVM), Gradient Boosting (XG Boost), Random Forest (RF), Decision Tree (DT), Logistic Regression (LR), K Nearest Neighbors (KNNs), and Artificial Neural Network (ANN), were efficiently employed in this study for the identification of initial signs of BS. The suggested ANN has a recognition accuracy of 94.35%, whereas RF has an identification rate of 94.15%. Both have about identical forecast accuracy for BS. The findings of the study revealed that ML and DL approaches have the potential to improve the identification of a variety of illnesses, such as BS, hence reducing the load and subjectivity issues in the medical field that existed owing to earlier traditional methods.

Keywords: Brain Stroke, Machine Learning, Deep Learning, smart diagnosis, Early Symptom Predictions

Introduction:

The survival of humans depends on the efficient operation of multiple body systems. BS is a severe clinical condition that can be life-threatening and typically occurs around the age of 65. Strokes damage the brain like how heart attacks affect the heart. Stroke causes localized or, in some cases, widespread impairment of brain function that lasts for more than 24 hours and can potentially result in death, with no apparent cause other than its origin in the brain. [1]. It occurs due to either a lack of circulation to the brain or a rupture and bleeds of cerebral blood veins. When a rupture or blockage happens, both oxygen and blood cannot reach cerebral cells. It has become the sixth most common cause of mortality across both advanced and developing nations [2]. Every year, about sixteen million individuals globally suffer from strokes, which have serious socioeconomic effects. Furthermore, the cost of medical therapy is rising, necessitating the development of advanced technologies that can help with clinical evaluation, care, identifying medical events, providing suitable therapies, attempts to recover, and [3][4].

Prompt medical intervention significantly improves a stroke patient's chances of full recovery. Immediate medical attention is crucial, as delays can result in permanent disability, severe brain damage, or mortality. According to the National Heart, Lung, and Blood Institute, strokes can be caused by various factors, including poor nutrition, physical inactivity, substance or tobacco use, genetic predisposition, and underlying medical conditions [5]. A stroke can impair memory, mobility, sight, language, and mental abilities [6]. Recognizing the symptoms of a fatal stroke can be difficult and time-consuming for medical professionals. The success of treatment and duration is mostly determined by the degree of severity of signs and the level of damage to organs. Early detection and management are critical in reducing long-term effects on an individual's general well-being; therefore, both medical practitioners and patients require an accurate and timely diagnosis [7].

AI and ML breakthroughs have resulted in major improvements in medical and automated clinical procedures. ML is a cornerstone of modern technology, and it plays a crucial role in the early detection of variant illnesses. In the healthcare industry, ML is critical for identifying and anticipating illnesses such as BS and coronary artery conditions [8][9]. Using demographic data, ML and DL algorithms can detect individuals at a higher risk of developing illnesses. These approaches can forecast the likelihood of illness progression by examining factors like sugar levels, blood pressure, age, and clinical history. Demographic characteristics are typically a collection of parameters obtained from quantitative and text data that is provided in the form of a table relevant to the topic under examination. Earlier studies used machine learning algorithms to forecast initial signs of illnesses through demographic information. However, in the past decade, DL algorithms have been deemed more reliable and generalized in predicting early signs of illnesses utilizing demographic information, depending on the amount of data used for the training set [10].

The classification of early signs of BS primarily relies on traditional diagnostic methods, which face challenges such as inaccurate identification, subjectivity, missed subtle differences, and prolonged diagnostic time [11]. ML and DL-based evaluation is a more effective alternative, offering greater accuracy, eliminating subjectivity, detecting subtle differences, and delivering faster results. Modern approaches to BS diagnosis rely significantly on innovative algorithms, and scholars have been experimenting with new machine learning and deep learning architectures to improve illness identification rates and generalizations [12]. Introducing new tools into the medical field presents a huge advancement that opens up new possibilities for enhanced accuracy, effectiveness, and early identification of multiple illnesses, including brain stroke. The utilization of smart tools for improvement in healthcare has proven to be an encouraging move towards precise diagnosis. However, differentiating the initial signs of distinct brain illnesses can be difficult, requiring technical skills and topic knowledge [13].

Previous machine learning research has primarily focused on predicting cardiovascular events, with relatively limited studies on brain stroke prediction. The primary purpose of this study is to demonstrate how AI-powered systems can forecast the onset of BS. The study tested different approaches on a freely accessible training set and compared them to determine the best approach to BS recognition. There is a reasonable amount of previous work on BS, with most researchers relying on ML models. This study incorporates a DL approach alongside ML models to explore DL's role in BS prediction, paving the way for future research in this domain. DL models are considered more accurate and generalized than ML approaches, making them more reliable for real-world deployment. The main contributions of this research are given below.

1. A large set of demographic data was collected related to BS and performed feature engineering under the supervision of health experts.
2. We utilized various preprocessing techniques including data balancing to enhance the training procedure.
3. Multiple prediction approaches, like DT, SVM, RF, KNN, and Gradient Boosting (XG Boost) were trained and tested.
4. A DL model was also trained and fine-tuned to predict the early symptoms of BS.

Objectives:

This study aims to improve early symptom prediction of BS via advanced AI techniques and tackle the class imbalance issue. The research consists of a comparative evaluation of multiple ML approaches like RF, LR, DT, SVM, XGBoost, and the DL approach ANN. The in-depth evaluation of these approaches helped the authors discover the approach that successfully predicts BS symptoms. The analysis was performed by utilizing evaluation metrics like accuracy, recall, precision, F1-score, PR, and ROC curves. Moreover, the outcomes of these algorithms have been depicted for an improved perception of algorithm performance. With the application of state-of-the-art AI approaches, this study aims to minimize the occurrence of BS across the globe by enabling quick diagnosis.

Related work:

Over the decades, researchers have made significant advancements in brain stroke recognition and risk assessment using ML and DL techniques. This section focuses on a number of the most recent developments and novel techniques in this field. The authors of [14] used an upgraded RF method to identify BS. This method was used to evaluate the risk indicators linked with strokes and allegedly surpassed previous SOTA methodologies. However, this research focuses only on specific types of BS and may not be helpful to emerging variations of BS in the future. Research carried out in [15] revealed the invention of an ML technique for forecasting heart attacks. The method was developed using various methods such as Naïve Bayes (NB), SVM, and DT, and the outcome was contrasted. The maximum precision obtained was 60%, which is considered quite low.

To examine the association between a feature and its impact on brain stroke, a method based on regression was used in a study [16]. The authors of that study used the DT technique and the machine learning algorithm KNN to forecast brain strokes. The study found that the DT method was far better at predicting the likelihood of BS. The authors in [17]18] employed probability-adaptive and neighboring KNN approaches to detect liver illness, utilizing an uneven training set as an emphasis. Various approaches can be used to address these challenges, including data processing through visual information, quantitative data analysis, and the integration of mixed inputs. In [18], the researchers employed directed learning to identify long-term risk factors for heart diseases. This research utilized multiple ML techniques, such as LR, RF, and SVM, and assessed their performance utilizing metrics such as accuracy, recall, and F1-score. An evaluation approach known as k-fold was used to test these methods, and the research concluded that LR outperformed the other alternatives with an overall accuracy of 72.06 %.

Kansa D et al. [19] examined the risk of stroke utilizing DT, ANNs, and NB for evaluation of results. Their approach intended to identify strokes using the data presented. The authors evaluated the algorithms' accuracy and AUC. They compared these metrics to determine the efficiency of every approach in forecasting stroke incidences. The research revealed the advantages and limitations of every method in terms of accuracy and efficiency in stroke forecasting. The research in [20] used ML techniques and found an ensemble strategy performed with 88% precision, 88.3% recall, and 87.8% accuracy. This study used the synthetic minority oversampling approach using cross-validation techniques with ten variations of K, emphasizing the reliability of balanced datasets along with information preprocessing for constructing improved machine learning models. Liu et al. [21] predicted the clinical result of a BS utilizing imbalanced and partial physiological parameters. Before classification, missing attributes were addressed using RF regression. Subsequently, a DL-based automated hyperparameter tuning approach was applied to an imbalanced training set to predict BS. Trials were conducted using a training collection that included 43,400 medical records and 783 instances of BS. When comparing the forecasting approach with various previous methods, the number of misclassification rates was reduced by an overall percentage of 51.50%, to around 19.1%. According to the reported outcomes, the obtained false positive rate, efficiency rate, and specificity are 33.12%, 72.00 %, and 67.39%, respectively. Similarly, Abdel-Fattah et al. [22] suggested a new ML strategy for Chronic Kidney Disease (CKD) diagnosis based on Apache Spark. The authors conducted studies utilizing all attribute sets and attributes picked using Relief-F. The results revealed that ML techniques such as SVM, GBM, and DT achieved a high detection rate of 100%.

Venkatesan et al. [23] proposed an AI-based technique for quick and accurate detection of CKD. During the phase of preprocessing, the researchers employed a class imbalance technique and dealt with null values using the KNN imputer. The findings indicated that the XGBoost machine learning method achieved a high detection rate of 98% in early-stage CKD detection. Azam et al. [21] investigated and contrasted 3 methods, namely LR, DT, and RF. They used a variety of methods for data preprocessing and balanced the data. The outcomes from the research show that the RF model significantly outperformed both machine learning designs, with an average mean accuracy of 99.89%. Over the last 10 years, BS has emerged as the leading cause of mortality globally. Scientific scholars have developed a variety of data-interpreting approaches to assist medical professionals in classifying cardiac problems. KNN is an excellent data analysis method to tackle categorization challenges.

Showman, M. et al. [24] used KNN to automatically diagnose cardiac problems. The results showed that KNN outperformed both the ANN and the ensemble approach. It was subsequently noticed that because of the type of training data, KNN can only be used for narrowly specified classes. This suggests that while KNN may be effective for detecting specific heart conditions, it may not be suitable for more intricate feature data. Despite this drawback, the researchers believe that KNN can be a valuable tool in the diagnosis of heart disorders, particularly when data is sparse or specialized. The researchers also indicate that KNN can be combined with some other approaches to enhance precision. The study in [25] created an automated method for quick diagnosis of cerebral strokes. The stacking strategy used in this work is new and performs admirably. The results indicated that the proposed mounting strategy had the highest rate of success at 97.4%. Scientific scholars have developed a variety of data interpretation approaches to assist medical professionals in classifying cardiac problems. As a result, using novel data identification algorithms may require fewer physical tests. A rapid and effective detection tool is required to reduce the number of mortality related to cardiovascular illnesses and brain stroke.

Patel J. et al. [26] used multiple ML variations, such as the DT classification algorithm, to find the most effective results in heart disease detection using the free and powerful framework WEKA. The methods used were J48, LR, and RF. The outcomes exhibited that the J48 method, a type of DT, achieved a higher recognition rate, compared to the other methods. The research in [27] used ML methods to detect CKD. The preprocessing phase involves interacting with incomplete values for attributes and selecting appropriate attributes. The study employed a total of eleven machine learning techniques, including regression, tree-based, and probabilistic methods, with an impressive 100% accuracy on a California University, Irvine training set. AdaBoost, ETC, DT, RF, and XG Boost, and RF demonstrated good performance in terms of accuracy. Singh et al. [28] tested different strategies for diagnosing strokes utilizing a DL model built on a public dataset of cardiovascular abnormalities. The study created a highly thorough diagnosis strategy for stroke abnormalities, reaching an accuracy rate of 99.07% following a detailed evaluation and incorporation of forecasting skills using various methods and varied techniques.

Proposed Method:

This study aims to improve the early detection of BS utilizing AI-based technologies to enhance diagnostic speed and accuracy. This section outlines the procedures and resources used to diagnose BS early indications. It describes the training set and the approaches to ML and DL. Figure 1 depicts the basic structure of the proposed BS prediction technique.

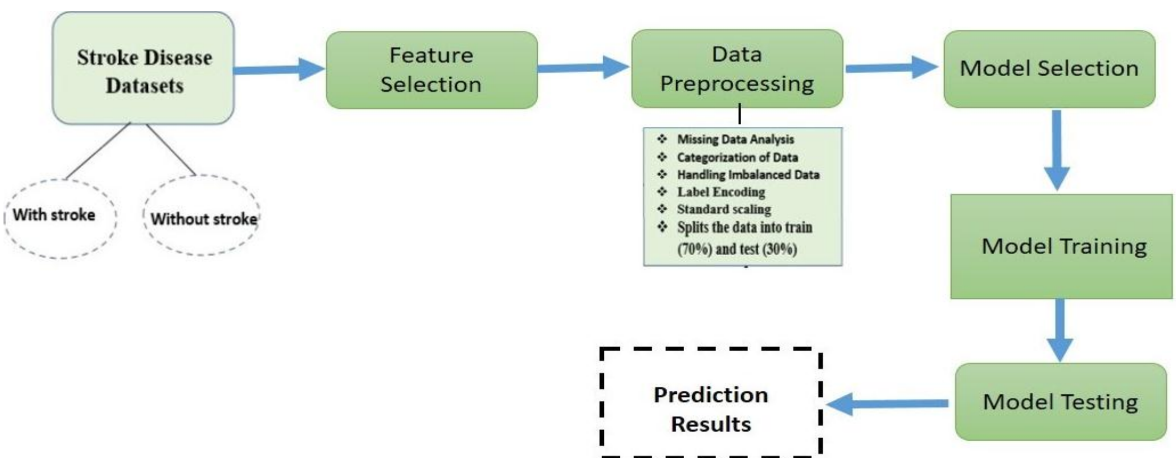

Figure 1: The basic steps of the Study

Training Dataset:

Supervised learning relies heavily on the training set. It serves as the foundation for developing an AI-based model and significantly impacts subsequent processes. The accuracy and generalization of a technique are heavily dependent upon the characteristics of the data set used to create and train the algorithm. For this research, we used a freely accessible dataset [29]. The training dataset contains detailed information on hypertension, age, BMI, smoking, and heart disease, as shown in Table 1. The dataset, which includes 12 medical parameters, allows for the prediction of individuals at risk of experiencing BS. The dataset is separated into testing and training sets, containing 12 columns and 5110 rows, with each row representing a single patient record. The original training set is considered imbalanced because it contains 4861 normal records and 249 abnormal records, which need to be balanced for improved and generalized model outcomes. This dataset is an invaluable resource for researchers seeking to understand and forecast the incidence and risk factors related to BS. Figure 2 depicts the utilized features with correlation among them. Table 1 shows the details of the training set.

Data Preprocessing:

ML and DL use a variety of strategies to purify and standardize text information, including reducing word forms, deleting extraneous components, and assuring consistency. These processes enhance data quality, allowing algorithms to focus on relevant patterns and features. Several data pre-processing strategies, including missing information, dealing with unbalanced information, encoding, scaling of features, and data portioning, were used in this study.

Missing Attributes Handling: Incomplete data is a typical difficulty in pattern understanding, and it can be influenced by some reasons, including mistakes in data collection or saving, incomplete interviews, or gaps in the actual data obtained [21]. In this research, the training set was examined to compensate for every instance of missing data. The other characteristics with no empty records except the BMI had 201 null values. The KNN imputer was used to substitute the values that were missing. The KNN Imputer serves as a sci-kit-learn information preparation approach that replaces missing values for attributes in the dataset. The current study employs the KNN technique to investigate values that are absent using the data points of the nearest entries in the appropriate attribute field. First, the dataset was scanned to identify missing values in various attributes. Then, for each missing value, the algorithm was used to find **K** nearest neighbors, data points without missing values, using a distance metric such as Euclidean distance. These neighbors were selected from the same feature space to maintain consistency and accuracy in the imputation process. Finally, the missing values were replaced by the weighted average (for numbers) or the most frequent value (for categories) of the selected neighbors. Closer neighbors had a greater influence, meaning their values carry more weight in determining the imputed result.

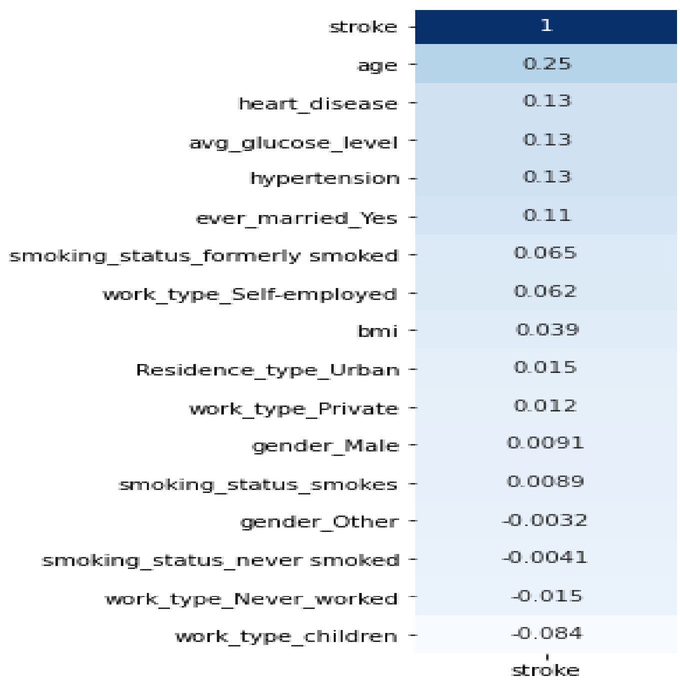

Figure 2: Shows the utilized features and their correlation

Table 1: Shows the details of training attributes

S N0.	Attribute	Description
1	Id	Unique identifier
2	Gender	"Female", "Male", or "Other"
3	Age	Age of the individual
4	Hypertension	0 = No hypertension, 1 = Has hypertension
5	Heart disease	0 = No heart disease, 1 = Has heart disease
6	Ever married	"No" or "Yes"
7	Work type	"children", "Govt job", "Never worked", "Private", "Self-employed"
8	Residence type	"Urban", "Rural" or

9	Avg glucose level	Average blood glucose level
10	Bmi	Body Mass Index
11	Smoking	"smokes", "Unknown", "never smoked", "formerly smoked",
12	Stroke	Had a stroke=1, No stroke=0

Rescaling Features Values: Rescaling in machine learning is a preprocessing technique that transforms non-normally distributed continuous data into a normal distribution, typically with a mean of 0 and a standard deviation of 1. In most procedures, the degree of scaling of the features influences both the method's efficiency and convergence time; hence, scaling is critical. The standard scalar package Sk-learn was used to rescale features in this investigation.

Feature Selection: Feature selection is the process of selecting the most significant qualities, which are the major factors that influence the target anomaly while eliminating those that have little to no impact on disease diagnosis. In the suggested task, only the person's ID was removed from the attribute list.

Attribute Encoding: Feature or label encoding in ML is a technique for translating category data into numerical form, allowing models to process these characteristics. There are different approaches, each with its own set of advantages and disadvantages, but one must be chosen based on the nature of the assignment. These methods are critical for dealing with qualitative features in ML techniques. In this method, label encoding was used to transform all qualitative attributes into numeric values.

Dataset Splitting: Dataset splitting is the process of splitting training data into subgroups to train and validate the suggested method. After finishing the training stage, the model is helped with an additional set of data known as the testing set. The data for training is used to educate the algorithm, while the set used for validation helps optimize parameters and avoid overfitting. A popular technique is to set aside 20% for validation and 80% for training. Proper partitioning guarantees that the model can generalize successfully, which is critical for achieving consistent performance on fresh data.

Dataset Balancing: While using an uneven training dataset to train an ML system may improve accuracy, other metrics such as recall and precision are insufficient. Unbalanced data, if not handled properly, will produce erroneous conclusions and forecasts. As a result, these uneven features must be addressed first to create a more effective system. As a result, an oversampling strategy known as SMOTE was used to build and maintain balance between the training classes. SMOTE (Synthetic Minority Over-sampling Technique) is a machine learning strategy for dealing with category imbalance that involves developing synthetic instances for the minority category. It generates new characteristics by interpolating between the existing minority instances, allowing the model to gain knowledge more efficiently from marginalized classes while avoiding bias favoring majority categories.

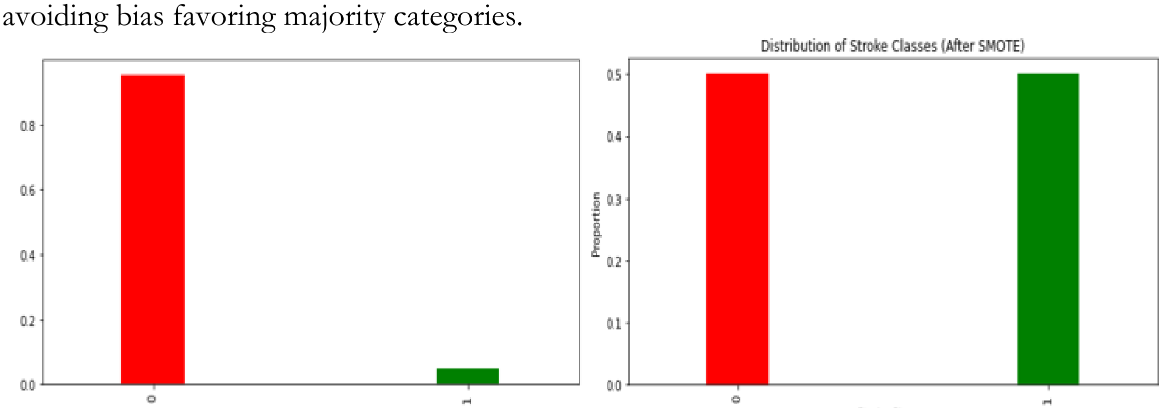

Figure 3: (a) Original Training dat (b) Upsampling Training Data via SMOTE

Machine Learning (ML):

ML is a branch of AI that enables algorithms to learn from incoming feature sets and make autonomous decisions for specific tasks without requiring explicit programming. These techniques enhance their performance over time by detecting patterns and associations within the data and can be adapted to integrate new, previously unknown information. This approach enables systems to improve their accuracy and efficiency in addressing complicated issues in a variety of disciplines, such as smart illness sign recognition. In the proposed study, various ML approaches, including LR, SVM, and RF, are employed, as outlined below.

LR is an estimation approach used mainly for binary classification issues in which the result is either one or zero. It uses logistic mapping to map the link between independent variables and probability values for a certain result, producing values that range from 0 to 1. LR, unlike linear regression, is best suited for categorical dependent variables, making it widely applied in areas such as health, finance, and social sciences for problems such as disease detection or customer sentiment analysis. The approach is straightforward but effective, frequently acting as a basis for contrasting with deeper systems. XGBOOST is a well-known, freely available ML package that is utilized for tasks like regression, classification, and rating. It is a robust and optimized package designed to be extremely compact, lightweight, and resilient. XGBoost makes forecasts with a combination of decision trees. During training, it generates a chain of DTs in which each tree attempts to correct the errors made by the previous one. This method enables XGBoost to deal with non-linear and complicated associations in feature collection, which can mostly produce higher accuracy than conventional machine learning techniques. XGBoost is commonly utilized in industries and has proven useful for several data science applications.

DT is a widely used approach for data interpretation and analysis. This framework leverages a DT model for classification or prediction, facilitating informed decision-making based on a given dataset. KNN is an ML approach that uses clustering of similar feature values and data. KNN is a non-parametric technique, meaning it makes no assumptions about the underlying distribution of the input data. The KNN technique attempts to categorize an unknown input data point by comparing it with the K-nearest points of data in the source collection. The author of this study specifies K, which represents the number of nearest neighbors utilized in categorization. An Ensemble Learning (EL) approach like RF can help with categorization and regression problems. They operate by the pooled training of multiple DTs. When utilizing RFs to solve classification problems, the outputs of RF correspond to the categories forecasted by the majority of the DTs. EL, or a strategy, uses a variety of identities to address complex challenges and improve network performance. The RF models combine more than one sub-tree into their forecasts, as stated in the name, "gathering a huge number of decision trees on various small sets of input data and calculating an average to improve the predicted precision." To forecast the outcomes, the most crucial forecasts are used. The Random Forest outperformed the other ML models in terms of predicting the initial signs of a BS.

SVM is an ML algorithm that uses annotated inputs to categorize new data [24]. Judgment areas, or hyperplanes, are used to depict decision limitations [30]. A hyperspace is used to separate a group of feature data into multiple classes. A radial basis function filter with a parameter value set to one was applied to enhance feature extraction and improve model performance. SVM attempts to categorize incoming data by creating a set of rules that assigns each occurrence of an input quantity to the appropriate label or class, utilizing the lowest amount of distortion and the highest (realistic) boundary. In this study, most machine learning models were trained using default settings. However, for the three models; XGBoost, Random Forest (RF), and SVM, we adjusted the parameter values to enhance their performance, as shown in Table 2.

Table 2: A list of parameters of ML Models

Model	Parameter	Value
XGBoost	N_estimators	200
	Learning rate	0.05
	Max_depth	06
	Sub_sample	0.8
RF	N_estimators	200
	Max_depth	20
	Min_Samples_split	5
	Min_Samples_leaf	2
SVM	Kernel	RBF
	C	1
	Gamma	0.1

Deep Learning (DL):

AI originated in the 1940s and has rapidly progressed over the years. Early research was built on traditional approaches like fuzzy logic, which provided the base for AI research. DL outperforms ML in complex tasks due to its ability to automatically grab useful information from raw data, eliminating the need for manual feature engineering. DL models, especially deep neural networks, excel at capturing complex patterns and dependencies, making them more effective for high-dimensional data like images and videos. Additionally, DL scales well with large datasets, improving performance as more data becomes available, whereas ML models can't capture the relationship between data and features in the case of large and complicated training datasets. DL models are known for robustness and generalization in real scenarios [31]. The proposed research employed a neural network comprising an initial layer aimed at getting input data and 6 fully connected layers followed by a final layer. The well-known optimization technique called Adam optimizer was used, with a Learning rate (Lr) of 0.0001 and an early stop mechanism was used to train the model up to a suitable number of epochs (the model stopped after 30 epochs automatically). Batch Normalization (BN) was used after each dense layer. Table 3 shows the structure of the ANN.

Table 3: Structural Info of the Proposed ANN

Layers	Neuron/ Value	Parameters	Value
Layers 01 and 02	64 x 2	Lr	0.001
BN		Batch Size	16
Drop Out	0.2	Epochs	100
Layers 03 and 04	128 x 2	Early Stop	30
BN		Activation	Relu
Drop Out	0.2	Optimization Technique	Adam
Layers 05 and 06	256 x 2	Loss Function	Binary_cross_entropy
BN		Training Set	0.80
Drop Out	0.2	Validation Set	0.20
Final Layer (Sigmoid)	1		

Training Environment:

Research trials utilizing ML and DL models require substantial resources, including high-performance hardware and advanced software. In this study, the primary hardware setup consisted of a Core i7 laptop with 16GB of RAM, a 500GB solid-state drive, and a multi-core processor. To develop and run the programs, a Python script containing several libraries such as Keras, TensorFlow, Seaborn, Scikit-learn, and Matplotlib was used, with Jupyter Notebook serving as the coding environment.

Performance Measurement:

To assess the effectiveness of an approach, its performance must be measured against previously unseen data using predefined evaluation metrics or criteria. The most often used criteria are Precision (P), Recall (R), Accuracy (A), and F1-score. Most of the current research uses these criteria to measure the performance of a model. The numerical representations of these formulas or metrics are provided below.

$$A = (\text{True Positive} + \text{True Negative}) / (\text{Total Predictions}) \quad (1)$$

$$P = \text{True Positive} / (\text{True Positive} + \text{False Positive} \quad (2)$$

$$R = \text{True Positive} / (\text{True Positive} + \text{False Negative}) \quad (3)$$

$$F_1\text{-Score} = 2 \times (P \times R) / (P + R) \quad (4)$$

Experiments and Results:

The goal of the study was to create an inexpensive AI-based system that allows for precise and prompt forecasting of BS's initial signs to treat the anomaly before it causes major consequences. The study used both DL and ML approaches, which are described more thoroughly below, along with validation results. Various approaches were used in this study to build an efficient approach that ranged from machine learning to deep learning models. This section presents ML and DL techniques, along with their test data performance. In this study, six machine learning algorithms, including DT, KNN, LR, XGBoost, RF, and SVM, were trained and tested on the suggested feature collection. All models demonstrated strong performance; however, RF outperformed the others, achieving a success rate of 94.20%. Table 4 shows the comprehensive results of all six machine learning algorithms and a model using ANN on the test set, together with the corresponding Receiver Operating Characteristic (ROC) and Precision-Recall (PR) curves.

Table 4: Outcomes of ML Approaches on Test Data

S No.	Approach	Accuracy (%)	Precision (%)	Recall (%)	F_1-Score (%)
1	DT	90.55	91.98	86.45	90.05
2	LR	92.80	93.20	90.05	91.60
3	KNN	91.65	93.05	87.28	90.08
4	XGBoost	93.66	94.30	90.15	92.05
5	SVM	93.80	95.35	91.30	92.45
6	RF	94.20	95.15	91.40	93.50
7	ANN	94.30	96.10	90.59	93.25

In addition to ML techniques, this study used multiple DL architectures. We continually modified the layer construction, regularization strategies, and width of the DL approach. After testing various deep learning architectures, the chosen framework consisted of an input layer, six Fully Connected (FC) layers, and a final layer with neuron configurations of 32, 64, 128×2, and 256×2, achieving an accuracy of 94.30%. Table 4 displays comprehensive outcomes for the suggested DL model.

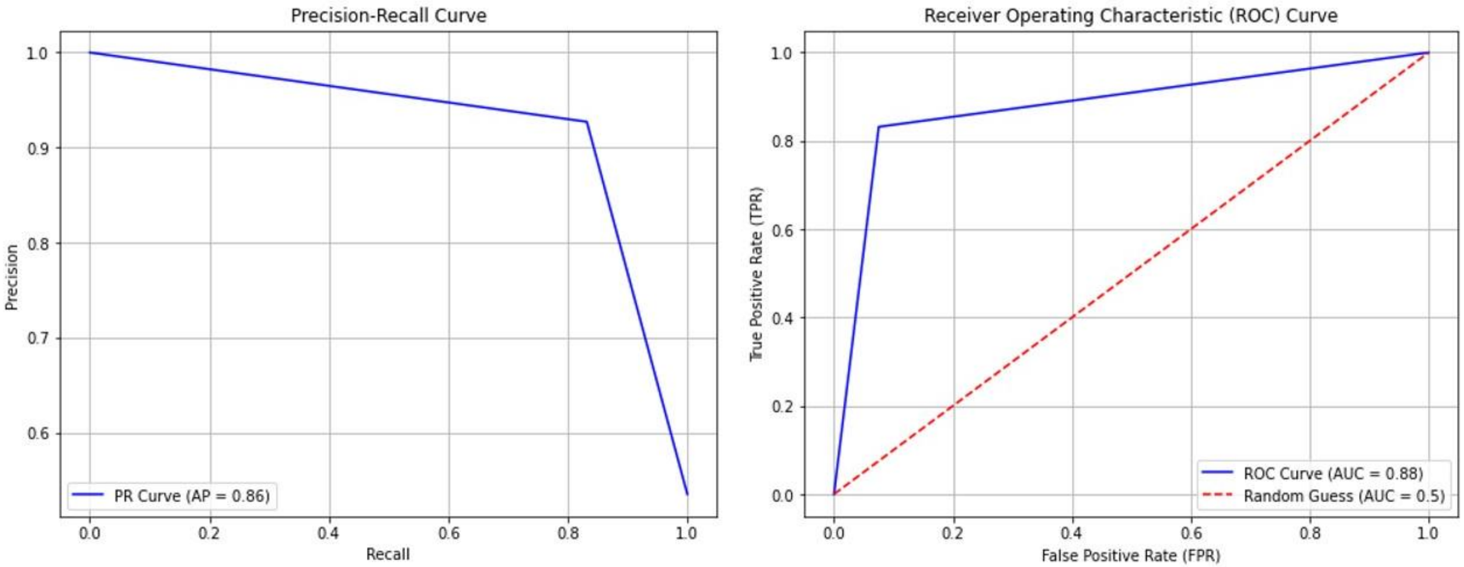

Figure 4: Shows DT Curves (a) PR (b) ROC

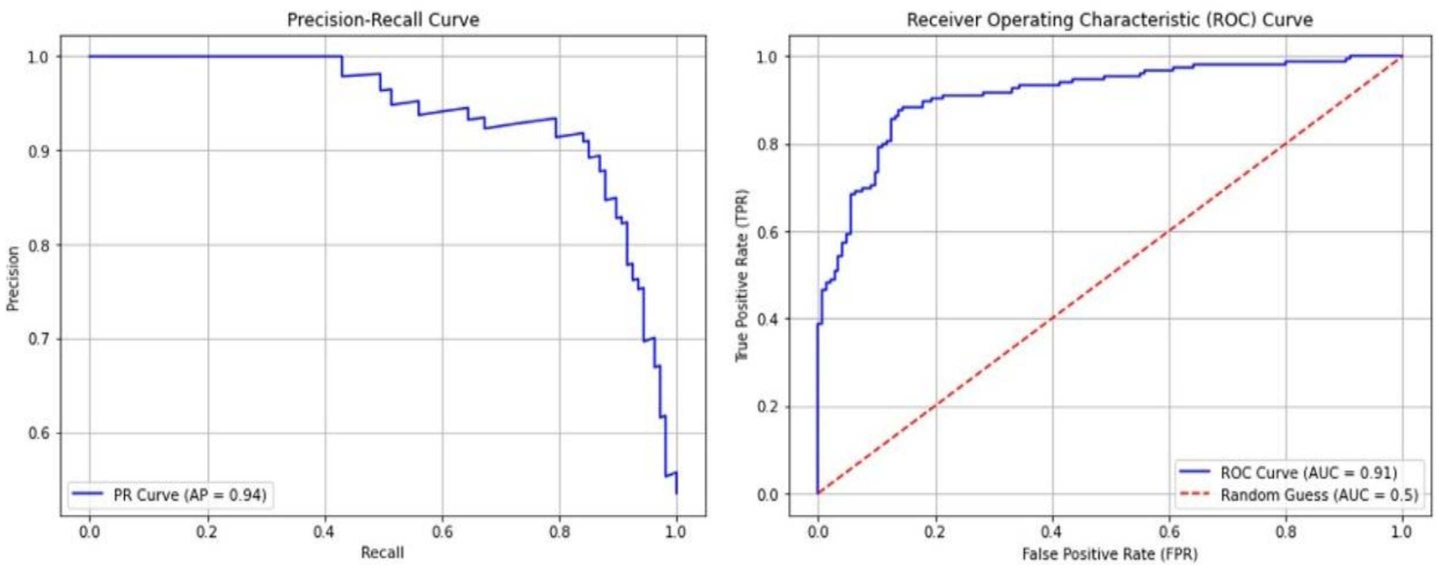

Figure 5: Shows LR Curves (a) PR (b) ROC

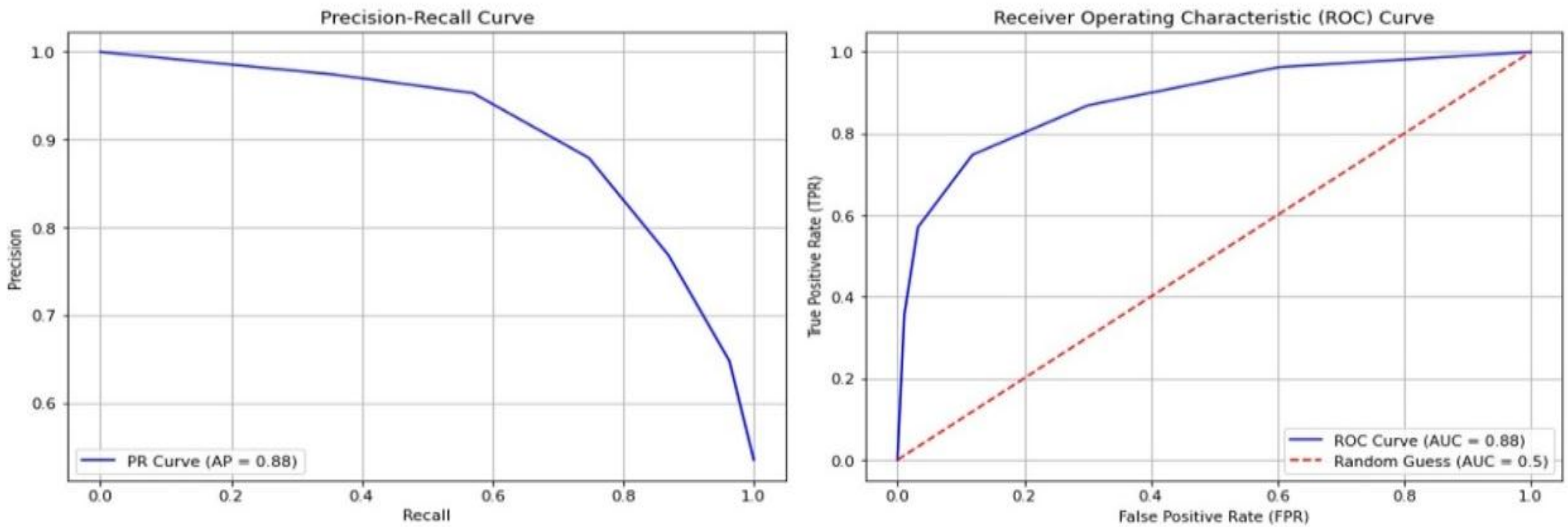

Figure 6: Shows KNN Curves (a) PR (b) ROC

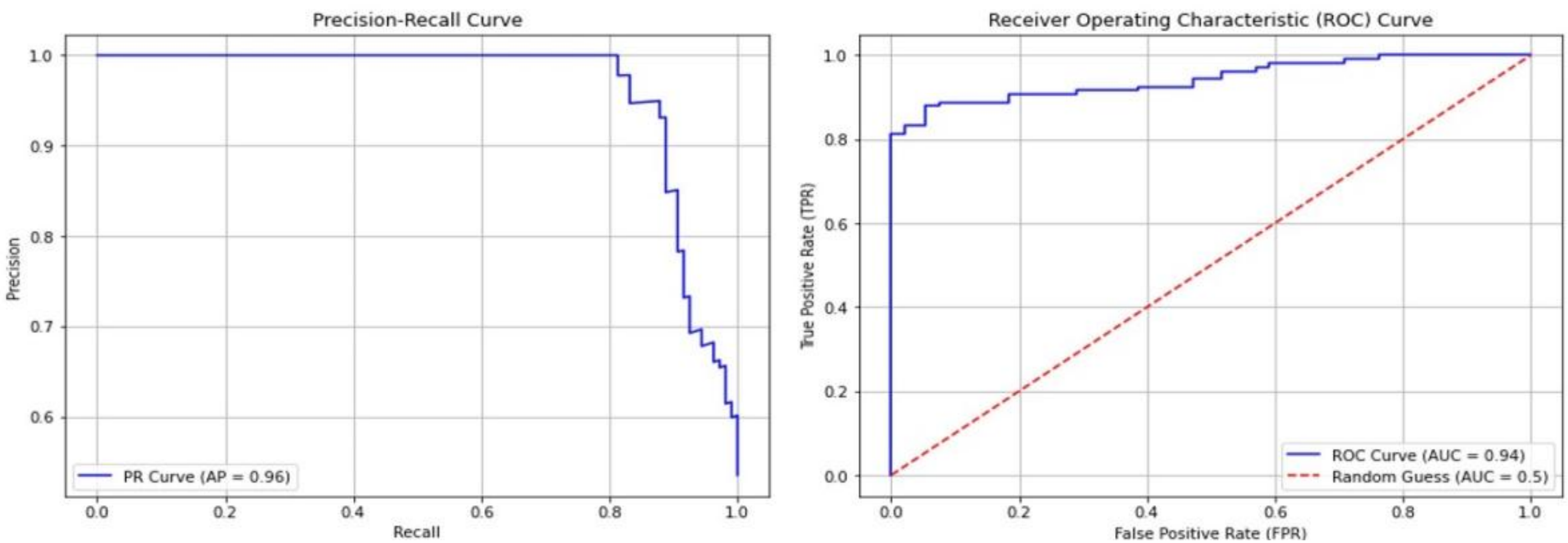

Figure 7: Shows XGBoost Curves (a) PR Curve(b) ROC

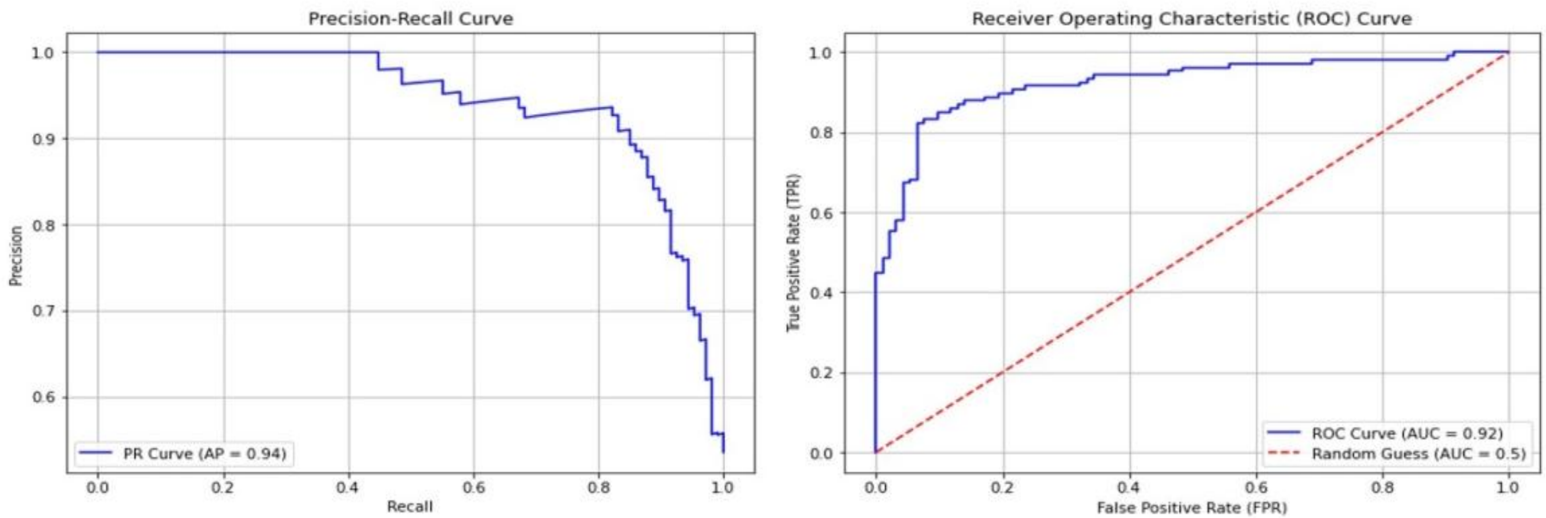

Figure 8: Shows SVM's Curves (a) PR (b) ROC Curve

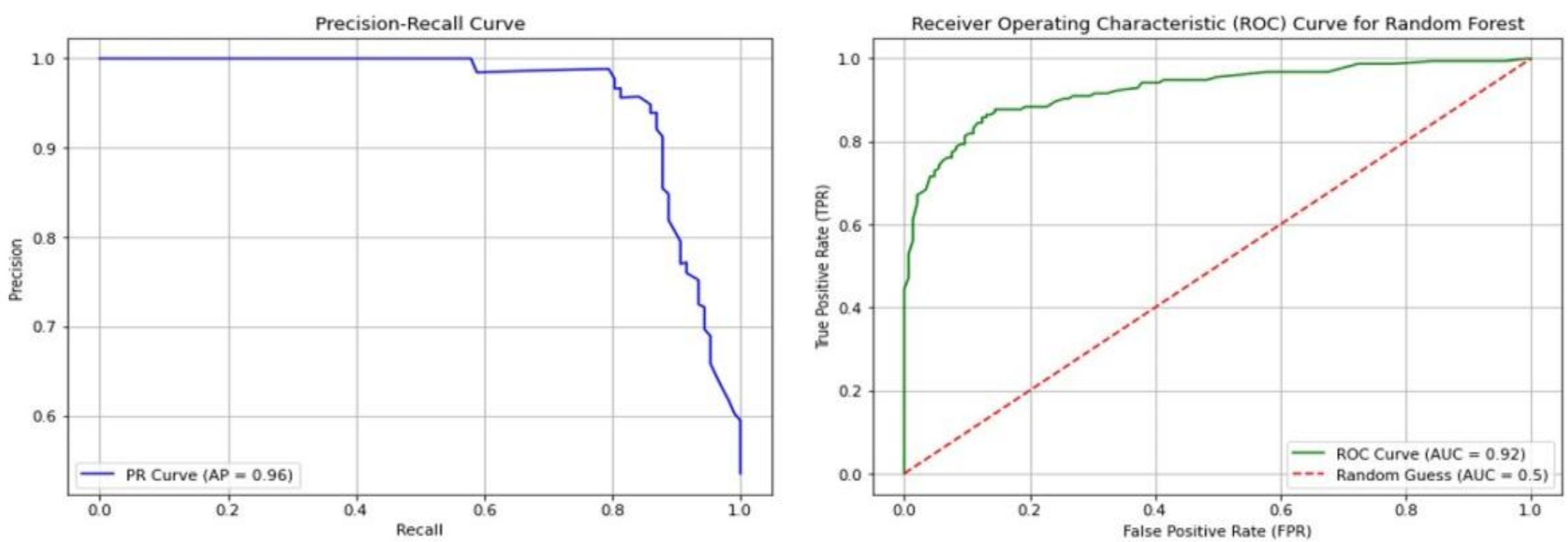

Figure 9: Shows RF's Curves (a) PR (b) ROC

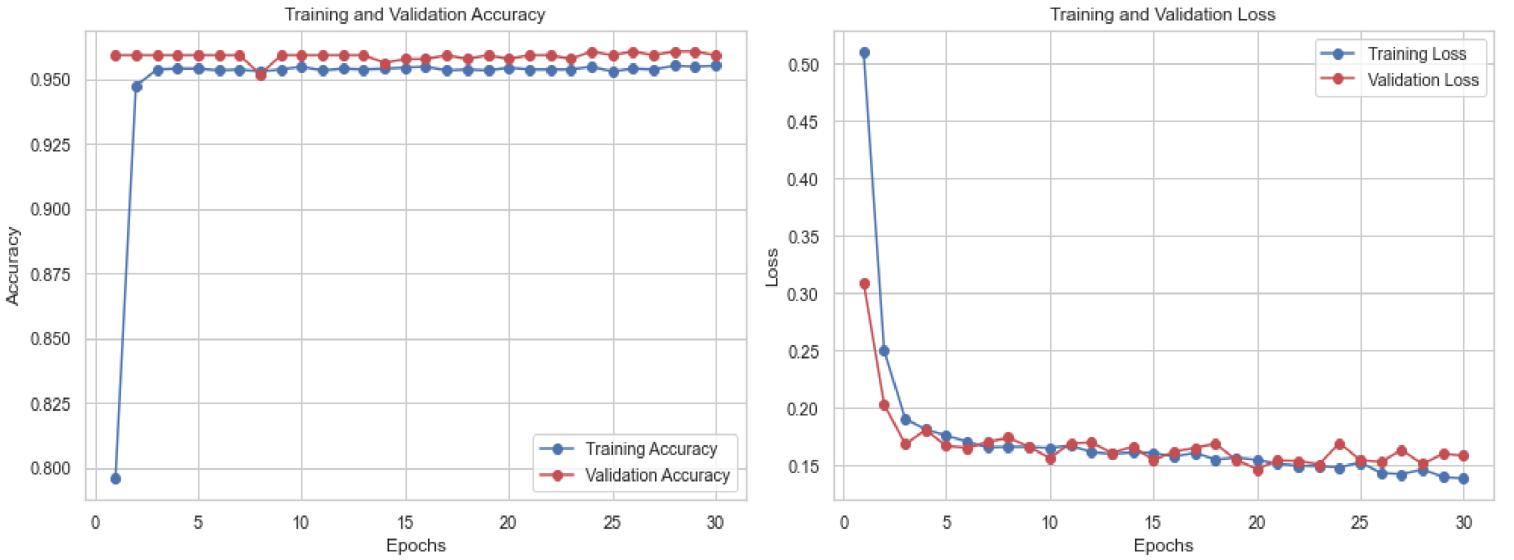

Figure 10. Training Progress of ANN (a) Accuracy (b) Loss

Performance Comparison of the Suggested Approach with Earlier SOTA Work:

To assess the reliability and accuracy of the proposed approach, we compared it with recent studies. Table 5 shoes a comparison of the proposed approach to past methodologies.

Table 5: Comparision of the proposed work with previous work

Ref#	Approaches Used	Accuracy
[32]	SVM, LR, and RF	83%
[33]	DT, SVM, NB, LR, KNN and RF	82%
[34]	XGBoost, KNN, LR, SVM and RF	92.32%
[35]	LR, AdaBoost, DT, RF, KNN and SVM	93.00%
[36]	DL (ANN)	92.39%
Proposed Work	SVM, RF, DT, KNN, LR and DL (ANN)	94.30%

Discussion:

In the proposed study, several ML models like SVM, RF, XGBoost, LR, DT, and KNN were trained to identify the early symptoms of BS. As compared to other ML models, RF performed well compared to other models with 94.20% accuracy. To optimize the RF model, default parameters were changed as shown in Table 2. Before training, the dataset was balanced using SMOTE to reduce bias toward the majority class, as the original dataset contained 4,861 normal cases and 249 BS cases. The main purpose of this study was to utilize and check DL-based models, known for more generalized results compared to ML models, and the ML models are used as a baseline approach. The suggested research tested a variety of DL structures like RNNs, 1D CNNs, ANNs, and LSTMs. In comparison to machine learning, DL-based techniques have been called data greedy; thus, the training set utilized in this research was of moderate size and was properly utilized while building DL systems. We iteratively updated the layer arrangement, involving basic parameters, but it was not suitable for the assignment at hand. As previously stated, these frameworks require additional data because of their complex structures. We applied a cost-effective ANN approach comprising an input layer, 6 hidden layers, and a final layer. The proposed DL model outperformed baseline ML models, achieving a 94.30% accuracy despite a moderately sized training set. Furthermore, expanding the training set is expected to further enhance performance. To achieve more solid, strong, and generalized outcomes, large amounts of training features must be collected, which will be explored in the future stage of this study.

Limitations of the Study:

The proposed study is based on the prediction of early signs of BS. To improve the generalization and accuracy of the approach, various ML models were used, and they produced good results. The study also utilized a DL-based approach with good results, but due to the imbalanced training set, it greatly affects the results of AI models specifically DL models. To overcome this limitation of the study, a large and balanced dataset is necessary, which can be collected from real scenarios, including health institutions, to further improve the accuracy and generalization of this approach, which would be useful for early diagnosis of BS to save human lives from mortality and impairment.

Conclusion and Future Work:

The purpose of this study was to focus on enhancing stroke possibility prediction by using multiple ML approaches as well as a DL technique called ANN. The study is based on a structured approach, which comprises data processing methods, model training, and result assessment and benchmarking. To mitigate model bias toward the majority class, the training dataset was balanced using the SMOTE strategy, ensuring improved representation of minority class samples. The generated results show the efficiency of the proposed approach, RF and ANN, with the highest prediction scores and ensure the use of AI techniques in early symptom identification of BS. Furthermore, selecting an appropriate ML or DL strategy depends heavily on the characteristics of the training data, which must be thoroughly analyzed before choosing a method. In the future, we aim to increase the size of the training dataset as well as to ensure class balancing, which provides a basis for more experiments with ML as well as with more deep models based on DL. The deep structure of a DL model guarantees more accuracy and generalization but also needs more data instances. The study also aims to use more advanced techniques based on DL and ensemble learning and utilize advanced optimization techniques to improve the overall prediction score of early signs of brain stroke.

References:

[1] Ayalew Zewdie, Finot Debebe, Aklilu Azazh, Sofia Kebede, Adam Laytin, Golnar Pashmforoosh, and Getaw Worku Hassen, “Prospective assessment of patients with stroke in Tikur Anbessa Specialised Hospital, Addis Ababa, Ethiopia,” *African J. Emerg. Med.*, vol. 8, no. 1, pp. 21–24, 2018, doi: https://doi.org/10.1016/j.afjem.2017.11.001.

[2] C. C. C. Andy Menke, Sarah Casagrande, “Cardiometabolic health in Asians with diabetes in the US,” *Diabetes Res. Clin. Pract.*, vol. 133, pp. 13–19, 2018, [Online]. Available: https://www.diabetesresearchclinicalpractice.com/article/S0168-8227(16)30580-0/fulltext

[3] O. A. K. B. Ovbiagele, L. B. Goldstein, R. T. Higashida, V. J. Howard, S. C. Johnston, “Forecasting the Future of Stroke in the United States: A Policy Statement From the American Heart Association and American Stroke Association,” *Stroke*, vol. 44, no. 8, 2013, doi: https://doi.org/10.1161/STR.0b013e31829734f2.

[4] S. Zia, A. N. Khan, K. S. Zaidi, and S. E. Ali, “Detection of Generalized Tonic Clonic Seizures and Falls in Unconstraint Environment Using Smartphone Accelerometer,” *IEEE Access*, vol. 9, pp. 39432–39443, 2021, doi: 10.1109/ACCESS.2021.3063765.

[5] G. L. Gary H, “National Heart, Lung and Blood Institute,” 2022, [Online]. Available: https://www.nhlbi.nih.gov/

[6] C. Elendu *et al.*, “Stroke and cognitive impairment: understanding the connection and managing symptoms,” *Ann. Med. Surg.*, vol. 85, no. 12, pp. 6057–6066, Dec. 2023, doi: 10.1097/MS9.0000000000001441.

[7] Lijing L. Yan , Chaoyun Li , Jie Chen,J. Jaime Miranda, Rong Luo, Janet Bettger, Yishan Zhu, Valery Feigin, Martin O’Donnell, and Y. W. Dong Zhao, “Prevention, management, and rehabilitation of stroke in low- and middle-income countries,” *eNeurologicalSci*, vol. 2, pp. 21–30, 2016, doi: https://doi.org/10.1016/j.ensci.2016.02.011.

[8] R. S. Jeena and S. Kumar, “Stroke prediction using SVM,” *2016 Int. Conf. Control Instrum. Commun. Comput. Technol. ICCICCT 2016*, pp. 600–602, Jul. 2017, doi: 10.1109/ICCICCT.2016.7988020.

[9] K. R. S. . Sabibullah Mohamed Hanifa, “Stroke Risk Prediction through Non-linear Support Vector Classification Models,” *Int. J. Adv. Res. Comput. Sci*, vol. 1, no. 3, 2010, doi: https://doi.org/10.26483/ijarcs.v1i3.60.

[10] A. E. H. Heba M. Afify, Kamel K. Mohammed, “Leveraging hybrid 1D-CNN and RNN approach for classification of brain cancer gene expression,” *Complex Intell. Syst.*, vol. 10, pp. 7605–7617, 2024, doi: https://doi.org/10.1007/s40747-024-01555-4.

[11] M. E. M. & T. G. Hee E. Kim, Alejandro Cosa-Linan, Nandhini Santhanam, Mahboubeh Jannesari, “Transfer learning for medical image classification: a literature review,” *BMC Med. Imaging*, vol. 22, no. 69, 2022, doi: https://doi.org/10.1186/s12880-022-00793-7.

[12] M. M. Taye, “Understanding of Machine Learning with Deep Learning: Architectures, Workflow, Applications and Future Directions,” *Computers*, vol. 12, no. 5, p. 91, 2023, doi: https://doi.org/10.3390/computers12050091.

[13] L. Pinto-Coelho, “How Artificial Intelligence Is Shaping Medical Imaging Technology: A Survey of Innovations and Applications,” *Bioengineering*, vol. 10, no. 12, p. 1435, 2023,

doi: https://doi.org/10.3390/bioengineering10121435.

[14] V. Bandi, D. Bhattacharyya, and D. Midhunchakkravarthy, "Prediction of brain stroke severity using machine learning," *Rev. d'Intelligence Artif.*, vol. 34, no. 6, pp. 753–761, Dec. 2020, doi: 10.18280/RIA.340609.

[15] F. S. Alotaibi, "Implementation of Machine Learning Model to Predict Heart Failure Disease," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 6, 2019, doi: 10.14569/IJACSA.2019.0100637.

[16] A. Khosla, Y. Cao, C. C. Y. Lin, H. K. Chiu, J. Hu, and H. Lee, "An integrated machine learning approach to stroke prediction," *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, pp. 183–191, 2010, doi: 10.1145/1835804.1835830/SUPPL_FILE/KDD2010_LEE_IMLA_01.MOV.

[17] P. Kumar and R. S. Thakur, "Diagnosis of Liver Disorder Using Fuzzy Adaptive and Neighbor Weighted K-NN Method for LFT Imbalanced Data," *6th IEEE Int. Conf. "Smart Struct. Syst. ICSSS 2019*, Mar. 2019, doi: 10.1109/ICSSS.2019.8882861.

[18] E. Dritsas, S. Alexiou, and K. Moustakas, "Cardiovascular Disease Risk Prediction with Supervised Machine Learning Techniques," *Int. Conf. Inf. Commun. Technol. Ageing Well e-Health, ICT4AWE - Proc.*, pp. 315–321, 2022, doi: 10.5220/0011088300003188.

[19] M. B. B. Selma Yahiya Adam, Adil Yousif, "Classification of Ischemic Stroke using Machine Learning Algorithms," *Int. J. Comput. Appl.*, vol. 149, no. 10, 2016, [Online]. Available: https://www.ijcaonline.org/archives/volume149/number10/adam-2016-ijca-911607.pdf

[20] M. T. Elias Dritsas, "Efficient Data-Driven Machine Learning Models for Cardiovascular Diseases Risk Prediction," *Sensors*, vol. 23, no. 3, p. 1161, 2023, doi: https://doi.org/10.3390/s23031161.

[21] 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, p. 101723, Nov. 2019, doi: 10.1016/J.ARTMED.2019.101723.

[22] and N. G. M. A. Abdel-Fattah, N. A. Othman, and N. Goher[25] M. A. Abdel-Fattah, N. A. Othman, "Predicting Chronic Kidney Disease Using Hybrid Machine Learning Based on Apache Spark," *Comput. Intell. Neurosci.*, vol. 15, pp. 1–12, 2022, doi: 10.1155/2022/9898831.

[23] and M. H. V. K. Venkatesan, M. T. Ramakrishna, I. Izonin, R. Tkachenko, "Efficient Data Preprocessing with Ensemble Machine Learning Technique for the Early Detection of Chronic Kidney Disease," *Appl. Sci.*, vol. 13, no. 5, p. 2885, 2023, doi: 10.3390/app13052885.

[24] R. S. Mai Shouman, Tim Turner, "Applying k-Nearest Neighbour in Diagnosing Heart Disease Patients," *IJIET*, vol. 2, no. 3, pp. 220–223, 2012, doi: 10.7763/IJIET.2012.V2.114.

[25] M. T. Elias Dritsas, "Stroke Risk Prediction with Machine Learning Techniques," *Maria Trigka*, vol. 22, no. 13, p. 4670, 2022, doi: https://doi.org/10.3390/s22134670.

[26] R. Deshmukh *et al.*, "Heart Disease Prediction Using Artificial Neural Network," *Int. J. Adv. Res. Comput. Commun. Eng.*, vol. 8, 2019, doi: 10.17148/IJARCCE.2019.8119.

[27] I. U. Ekanayake and D. Herath, "Chronic Kidney Disease Prediction Using Machine Learning Methods," *MERCon 2020 - 6th Int. Multidiscip. Moratuwa Eng. Res. Conf. Proc.*, pp. 260–265, Jul. 2020, doi: 10.1109/MERCON50084.2020.9185249.

[28] M. S. Singh and P. Choudhary, "Stroke prediction using artificial intelligence," *2017 8th Ind. Autom. Electromechanical Eng. Conf. IEMECON 2017*, pp. 158–161, Oct. 2017, doi: 10.1109/IEMECON.2017.8079581.

[29] "Stroke Prediction Dataset." Accessed: Feb. 15, 2025. [Online]. Available:

https://www.kaggle.com/datasets/fedesoriano/stroke-prediction-dataset
[30] E. Dritsas and M. Trigka, "Stroke Risk Prediction with Machine Learning Techniques," *Sensors (Basel)*., vol. 22, no. 13, p. 4670, Jul. 2022, doi: 10.3390/S22134670.
[31] G. F. Manakitsa, N., Maraslidis, G. S., Moysis, L., & Fragulis, "A Review of Machine Learning and Deep Learning for Object Detection, Semantic Segmentation, and Human Action Recognition in Machine and Robotic Vision," *Technologies*, vol. 12, no. 2, p. 15, 2024, doi: https://doi.org/10.3390/technologies12020015.
[32] M. M. Priyanka S R, "A Hybrid Feature Selection Method for Stroke Detection Using Machine Learning Approaches," *Int. J. Innov. Res. Sci. Eng. Technol.*, vol. 9, no. 7, 2020, [Online]. Available: https://www.ijirset.com/upload/2020/july/69_A_HYBRID_NC.PDF
[33] 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, doi: 10.14569/IJACSA.2021.0120662.
[34] Jose-A Tavares, "Stroke prediction through Data Science and Machine Learning Algorithms," 2021, doi: 10.13140/RG.2.2.33027.43040.
[35] L. Garcia-Temza, J. L. Risco-Martin, J. L. Ayala, G. R. Rosello, and J. M. Camarasaltas, "Comparison of different machine learning approaches to model stroke subtype classification and risk prediction," *2019 Spring Simul. Conf. SpringSim 2019*, Apr. 2019, doi: 10.23919/SPRINGSIM.2019.8732846.
[36] S. Rahman, M. Hasan, and A. K. Sarkar, "Prediction of Brain Stroke using Machine Learning Algorithms and Deep Neural Network Techniques," *Eur. J. Electr. Eng. Comput. Sci.*, vol. 7, no. 1, pp. 23–30, Jan. 2023, doi: 10.24018/EJECE.2023.7.1.483.