Journal of Applied Data Sciences
Vol. 6, No. 3, September 2025, pp. 2111-2126
ISSN 2723-6471
2111

Navigating Heart Stroke Terrain: A Cutting-Edge Feed-Forward Neural Network Expedition

S Phani Praveen[1], Jeevana Sujitha Mantena[2], Uddagiri Sirisha[3], Deshinta Arrova Dewi[4], Tri Basuki Kurniawan[5,*], Choo Wou Onn[6], Yorman[7]

[1]Department of CSE, PVP Siddhartha Institute of Technology, Kanuru, Vijayawada, Andhra Pradesh, India

[2]Department of CSE, SRKR Engineering College(A), Bhimavaram, Andhra Pradesh, India

[3]Department of Computer Science and Engineering, PVP Siddhartha Institute of Technology, Kanuru, Vijayawada, Andhra Pradesh, India

[4]Centre for Data Science and Sustainable Technologies, INTI International University, Nilai, Malaysia

[5]Post Graduate Program, Universitas Bina Darma, Palembang, Indonesia

[6]Faculty of Data Science and Information Technology, INTI International University, Nilai, Malaysia

[7]Universitas Bumigora, Mataram, Indonesia

(Received: December 28, 2024; Revised: February 3, 2025; Accepted: May 3, 2025; Available online: July 19, 2025)

Abstract

Heart stroke remains one of the leading causes of death worldwide, necessitating early and accurate prediction systems to enable timely medical intervention. While a variety of machine learning approaches have been employed to address this issue, including Logistic Regression, Decision Trees, Random Forests, Support Vector Machines, and K-Nearest Neighbors, these models often suffer from limitations such as overfitting, insufficient generalization, poor performance on imbalanced datasets, and inability to capture complex nonlinear patterns in clinical data. Additionally, many existing works do not comprehensively integrate both clinical and demographic features or lack rigorous evaluation metrics beyond accuracy alone. This study proposes a novel Feed-Forward Neural Network (FFNN) model for heart stroke prediction, designed to overcome the shortcomings of conventional models. Unlike shallow classifiers, the FFNN architecture employed here leverages multiple hidden layers and nonlinear activation functions to learn intricate relationships within the dataset. The dataset used comprises various attributes such as age, hypertension, heart disease, BMI, and smoking status, which were preprocessed through normalization, one-hot encoding, and imputation techniques to ensure data quality and model performance. Experiments were conducted using a stratified train-test split, and the model was trained using the Adam optimizer with carefully tuned hyperparameters. Comparative evaluations against baseline models (Logistic Regression, Random Forest, and SVM) were carried out using precision, recall, F1-score, and ROC-AUC as performance metrics. The proposed FFNN achieved the highest accuracy of 96.47%, along with substantial improvements in recall and F1-score, highlighting its superior capability in identifying potential stroke cases even in imbalanced datasets. This work bridges a significant gap in heart stroke prediction by demonstrating the effectiveness of deep learning models—specifically FFNNs—in extracting complex patterns from diverse patient data. It also sets the stage for further exploration of deep learning-based clinical decision support systems.

Keywords: Heart Stroke Prediction, Neural Network, Machine Learning, Predictive Modelling, Accuracy, Public Health, Cardiovascular Disease,

1. Introduction

A stroke, also known as a Cerebrovascular Accident (CVA), is a medical emergency in which the supply of oxygen and nutrients to parts of the brain is interrupted, leading to rapid brain cell death. Stroke is a critical condition requiring immediate medical attention [1]. Globally, stroke affects approximately 13 million people annually and results in over 5 million deaths. It is the second most common cause of adult disability worldwide and ranks fifth in the United States as a cause of mortality, while being the leading cause of long-term adult disability [2].

***Corresponding author: Tri Basuki Kurniawan (tribasukikurniawan@binadarma.ac.id)**
DOI: https://doi.org/10.47738/jads.v6i3.763

Stroke prediction refers to the identification of individuals at high risk of experiencing a stroke, based on factors such as age, sex, race, medical and family history, and lifestyle [3]. This field of research is essential because early identification of high-risk individuals can significantly reduce the burden of stroke. Effective prediction enables timely preventive measures and treatment, which greatly improves patient outcomes [4], [5].

However, stroke prediction faces several challenges. One of the key difficulties is the wide variety of potential causes of stroke, some of which are still not well understood. Moreover, risk levels vary significantly between individuals, and accurate models must distinguish between high-risk and low-risk populations to avoid unnecessary treatment and associated costs. Despite these issues, advancements in stroke prediction algorithms have led to more accurate identification methods in recent years [6].

The urgency of improving stroke prediction is evident when considering the broader impacts. After cancer and heart disease, stroke is the third leading cause of death worldwide. In the United States alone, stroke-related healthcare costs exceed $50 billion annually [7], [8]. Early detection and treatment can significantly reduce mortality and disability, alleviating both human suffering and economic burden [9]. Yet, model reliability remains a pressing concern. Some individuals suffer strokes despite having no known risk factors, while others with multiple risk factors may never experience one. Thus, predictive models must be both accurate and personalized. They must also effectively balance sensitivity and specificity to ensure clinical relevance and practicality [10].

To achieve this, large and diverse datasets must be leveraged, and prediction tools must be accessible to clinicians. Artificial intelligence (AI) has emerged as a promising solution, with machine learning algorithms capable of identifying complex patterns in medical data to enhance prediction accuracy [11]. Machine learning and data science provide tools for building robust stroke prediction models. These approaches allow researchers to process large volumes of patient data to identify risk-related features and train accurate classifiers. Such models often outperform traditional statistical techniques [12], [13], [14].

Given the complexity of stroke prediction, the limitations of classical machine learning models, and the presence of issues like class imbalance, we propose the use of a deep learning model. A feed-forward neural network built from scratch is particularly promising due to its ability to autonomously learn and capture intricate patterns in data. This model holds the potential to provide more reliable and accurate stroke prediction results [15], [16], [17].

2. Literature Review

Research on stroke prediction has evolved significantly in recent years, with various approaches aiming to identify individuals at risk based on clinical, demographic, and behavioral data. Several studies have explored machine learning and data mining techniques to enhance prediction accuracy and clinical relevance [18], [19], [20], [21]. In [22], researchers utilized the Cardiovascular Health Study (CHS) dataset to evaluate the performance of multiple machine learning algorithms. They proposed an integrated approach combining Decision Tree (DT), Principal Component Analysis (PCA), Artificial Neural Networks (ANN), and Support Vector Machine (SVM). This ensemble method yielded the best results among the tested configurations. However, a major limitation of this study was the constrained number of input parameters available in the CHS dataset, which may reduce the model's applicability in real-world scenarios with more complex patient profiles.

A novel direction was taken in [23], where the Disease Related Feature Selection (DRFS) technique was applied to social media data for stroke symptom detection. Natural Language Processing (NLP) methods were employed to extract relevant features from user-generated content. Although this method expanded the range of data sources for prediction, it introduced significant computational complexity due to the processing demands of unstructured text data, resulting in increased runtime and resource usage.

The work presented in [24] introduced a modified Random Forest (RF) algorithm designed specifically for stroke prediction. The enhanced RF model outperformed traditional algorithms in accuracy and robustness. Nevertheless, the study was limited in scope as it focused solely on ischemic stroke and did not explore model adaptability to other stroke subtypes or unseen data distributions.

Another comparative study evaluated DT, RF, and Multi-layer Perceptron (MLP) models for stroke prediction using standard datasets [25]. Accuracy levels ranged from 74% to 75%, with MLP demonstrating marginally superior performance. However, the study relied exclusively on accuracy as the evaluation metric, neglecting other important indicators such as precision, recall, F1-score, and AUC, which are essential for a balanced assessment in imbalanced medical datasets.

In [26], the performance of DT, Naive Bayes (NB), and SVM was analyzed, with the highest accuracy reaching only 60%. This relatively low performance underscores the need for more sophisticated models or richer datasets. By contrast, [27] reported a striking 95% accuracy using a combination of three classification algorithms—C4.5, JRip, and MLP. While impressive, the high accuracy came at the cost of increased model complexity and extended training and prediction time, which may hinder deployment in time-sensitive clinical settings.

Study [28] also focused on comparing the performance of NB, DT, and ANN in stroke prediction. DT again proved to be the most accurate among the three with a 75% accuracy rate. Although ANN offered potential advantages in capturing nonlinear relationships, it required more training time and tuning, highlighting a trade-off between performance and efficiency.

Feature selection remains a pivotal aspect of stroke prediction, and [29] proposed a new method to automatically select robust features from the CHS dataset. While the approach led to improved model generalization, its combination with SVM resulted in the generation of a large number of support vectors. This not only slowed down model inference but also reduced interpretability—an important factor in clinical decision-making.

The use of ANN with backpropagation algorithms has also shown promise in stroke prediction. Studies [30], [31], and [32] applied this approach to predict thromboembolic stroke with high accuracy. However, they encountered significant training challenges as the number of neurons and hidden layers increased. These models became computationally expensive and less practical without high-performance computing resources. Additionally, later follow-up studies [33], [34], [35] confirmed that while ANN models are powerful, their complexity must be managed to ensure scalability and clinical usability.

Across these studies, it is evident that no single model offers a perfect solution. Classical models like DT and RF are easier to interpret and deploy but may lack the predictive power of deep learning models. On the other hand, deep learning models like ANN and MLP provide higher accuracy but require extensive tuning and resources. These findings emphasize the importance of model selection based on context—balancing predictive performance with efficiency, interpretability, and scalability.

3. Material and Method

3.1. Proposed Methodology

The proposed methodology is presented in figure 1. The dataset is pre-processed to remove outliers and replace missing values. The pre-processed data is analyzed using exploratory data analysis to gain insights into the correlations between the various parameters. Feature importance is then used to select the most salient parameters that influence stroke risk prediction. Parameters with little or no influence on the target output are removed to reduce computational cost and the chance of model overfitting. A ratio of 80:20 is used for training and testing the model, with 80% being used to train and 20% to test. An accurate, precise, recall, and F1 score are used to determine whether the model is acceptable.

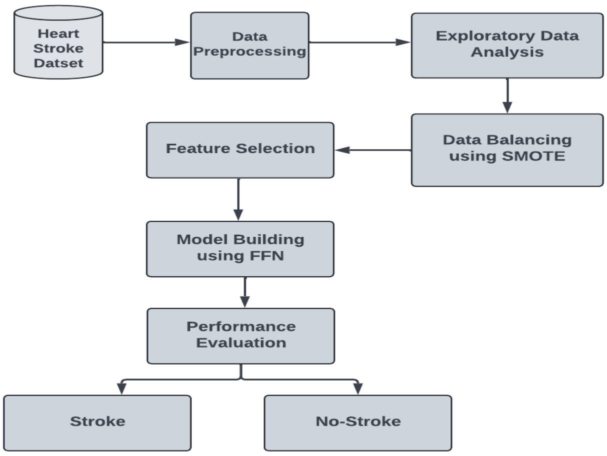


Figure 1. Proposed methodology

3.2. Dataset Description and Preprocessing

All Data for stroke prediction was collected from Kaggle [36]. The key information about the dataset is summarized in table 1. It is pertinent to mention that there are 201 missing entries in the BMI column, and roughly 30% of patients lack information regarding their smoking status, presenting a training challenge. In essence, the dataset is relatively small and exhibits a significant imbalance between stroke and non-stroke cases, posing difficulties for ML models in achieving accurate stroke predictions. Moreover, the presence of missing values further affected model accuracy. Notwithstanding these obstacles, the dataset remains viable for training ML models to forecast stroke. It's crucial, however, to acknowledge these dataset limitations and take corrective measures, including oversampling to address data imbalances and imputation techniques to handle missing values.

Table 1. Heart Stroke Dataset

Dataset Details	Stroke Prediction Dataset
Source	Kaggle [36]
Size	5110 rows, 12 columns
Columns	The ID, the gender, the age, the BMI, the hypertension, the heart disease, the marriage status, the level of glucose in the blood, the type of residence, the smoking status, and the stroke status are all included.
Target Variable	stroke' (binary): '0' (no stroke risk), '1' (potential stroke risk)
Class Imbalance	Significant class imbalance, '0' (4861 instances), '1' (only 249 instances)
Data Pre-processing	Applied to address class imbalance for improved predictive accuracy

The process is started with dataset pre-processing step to ensure data quality, enhancing the ML model performance. In this study, the data pre-processing, includes identifying outliers, handling null values, and reducing noise. Of the 5,110 total records, 201 had missing values for the BMI variable. To fix this, a Simple Imputer tool in Scikit-learn is used to substitute the median of the column for the missing data. Scikit-learn's Simple Imputer is a univariate imputation method that uses a suitable statistic (such as the mean, median, or most frequent) to replace missing values in each column. Additionally, the 'id' column is removed as it has minimal impact on the stroke risk. Furthermore, a single instance where the gender attribute was specified as 'Other,' was considered as an outlier and subsequently is removed.

3.3. Data Visualization

Data visualization is a crucial component of Exploratory Data Analysis (EDA), particularly in the context of stroke prediction. It involves the graphical representation of data through various visual tools such as bar charts, histograms, scatter plots, heatmaps, and box plots. The primary objective is to enable researchers and practitioners to observe patterns, trends, correlations, and outliers that may not be readily apparent from raw numerical data. In the domain of medical diagnosis, and specifically in stroke prediction, data visualization helps to assess the distribution of key variables such as age, gender, hypertension, heart disease, smoking status, and Body Mass Index (BMI). For example,

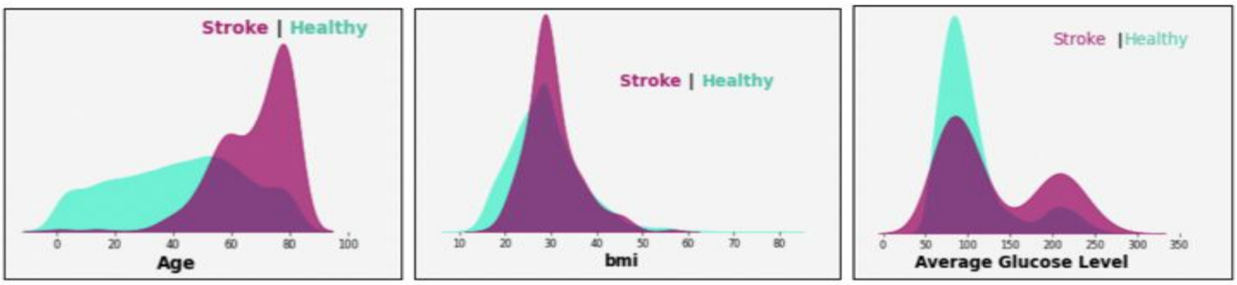
ISSN 2723-6471

a histogram can reveal whether age distribution is skewed, while box plots can highlight potential anomalies in blood glucose levels or BMI that may be associated with stroke risk. Such insights are essential for guiding data preprocessing decisions and informing feature selection for model development.

Moreover, correlation heatmaps are frequently employed to identify relationships between variables. For instance, a strong positive correlation between age and stroke incidence can inform the prioritization of age as a significant predictor in modeling efforts. Similarly, visualizing missing data patterns using heatmaps or matrix plots enables better handling of data imputation or exclusion strategies. Another key aspect of data visualization in this context is the comparison of stroke versus non-stroke populations across multiple features. Grouped bar charts and violin plots are often used to contrast feature distributions between these classes, which aids in understanding class imbalances and variable importance. This comparative analysis is critical for designing robust machine learning algorithms that can generalize well across different subsets of patient data.

3.3.1. Univariate-Analysis

Univariate-analysis is a statistical approach that centers on the examination of one variable or attribute at a time. Its primary objective is to comprehend the distribution, attributes, and characteristics of that specific variable independently. Univariate analysis serves as a tool for researchers and analysts to uncover patterns and behaviours within individual variables, excluding considerations of their interactions with other variables. In the context of strokes prediction, univariate analysis entails the scrutiny of individual variables to understand their distribution and characteristics, with a particular focus on their relevance to the target variable “stroke” as shown in figure 2.



Age Distribution by stroke | bmi Distribution by stroke | Glucose Distribution by stroke

Figure 2. Single-variable examination for heart stroke prediction

A BMI of 30 or above indicates obesity, and stroke risk is significantly elevated in obese individuals. The highest concentration of risk is evident within the BMI range of 25 to 30. Examining the graph in figure 3, it becomes apparent that stroke risk is notably elevated among individuals aged 60 to 80. Concerning glucose levels, the risk of stroke appears to be heightened across all levels.

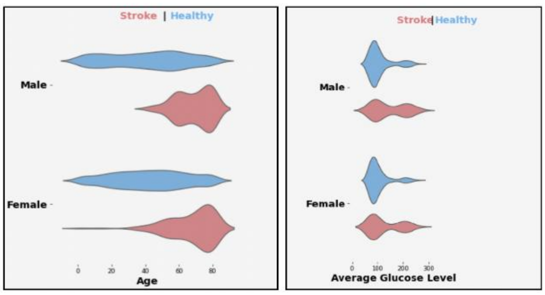


Figure 3. Age feature analysis in heart stroke prediction

Consequently, assessing stroke risk based solely on glucose levels may not be suitable. It is important to note that stroke risks are increased for individuals with glucose levels above 200. Both males and females have an increased risk of stroke with age, however the female population tends to have strokes earlier in life. Additionally, as glucose levels rise, both males and females face an elevated risk of stroke.

ISSN 2723-6471

Moreover, it can be observed that, there is a correlation between age and the likelihood of stroke. Additionally, there exists a notable association between stroke occurrence and both BMI and glucose levels. Consequently, these features are used in the modelling. It's worth noting that there is a need to normalize the average glucose level due to its substantial right-skewness. The Box-Cox normalization technique is used in this study for normalizing numerical features. Data distributions can be transformed to resemble normal distributions using the Box-Cox transformation [37]. Employing Box-Cox transformation can enhance the predictive capabilities of an analytical model by reducing the influence of random noise in the data. Furthermore, log transformation can be used as an alternative approach. One-dimensional exploration of heart stroke prognosis is shown in figure 4.

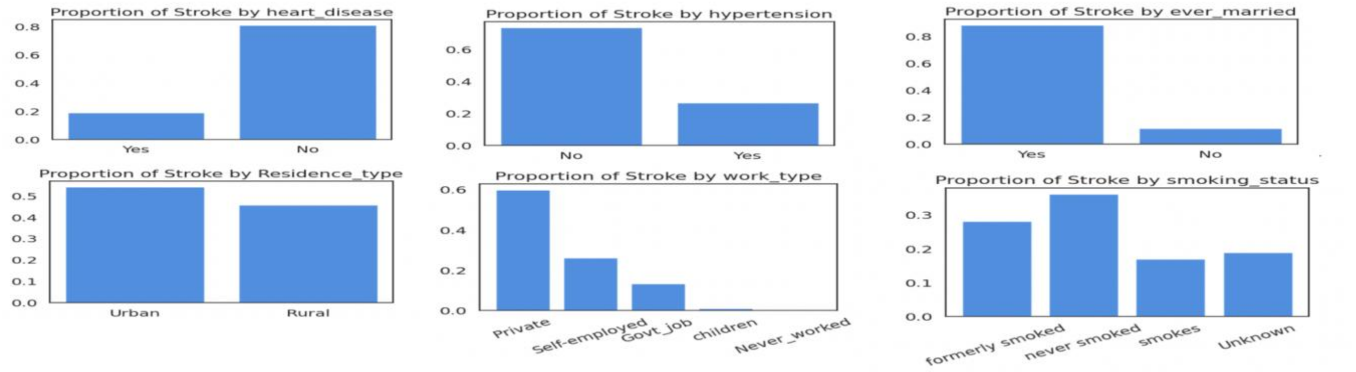


Figure 4. One-dimensional exploration of heart stroke prognosis

3.3.2. Bivariate Analysis

To understand the interaction between two variables and how they influence stroke risk, bivariate analysis is often performed (see figure 5).

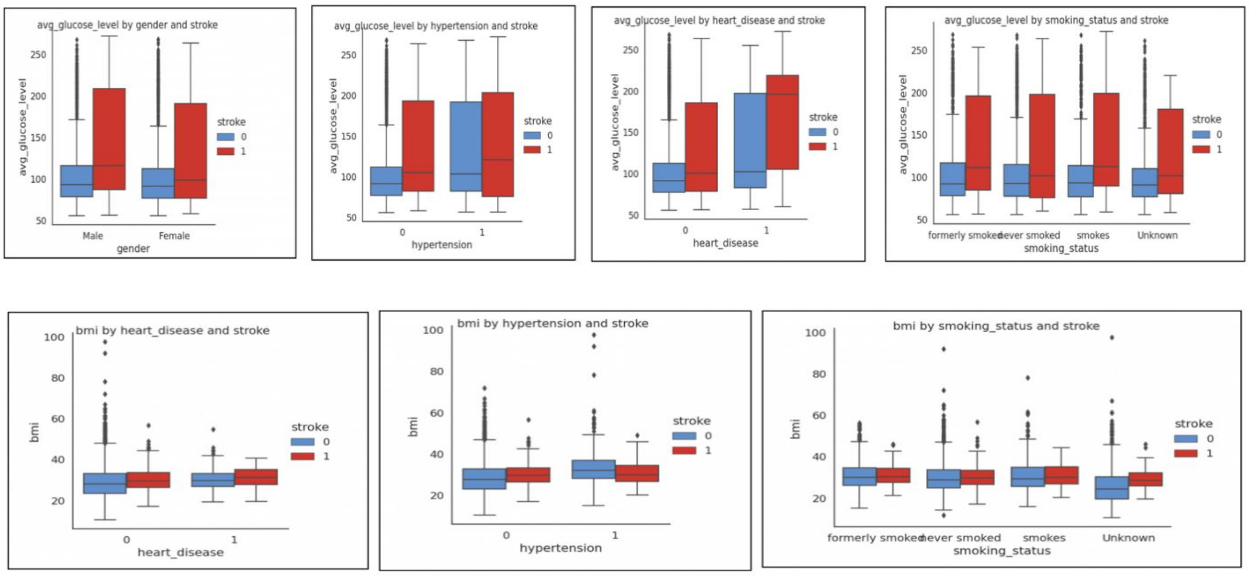


Figure 5. Two-dimensional exploration of various attributes for heart stroke prognosis

Figure 6 illustrated Pearson's correlation matrix for the bivariate analysis of the dataset. It is observed that both smoking_status and body bmi are inversely related to work_type, with BMI additionally showing a moderate age connection. In addition, late marriage in life is significantly associated with older age.

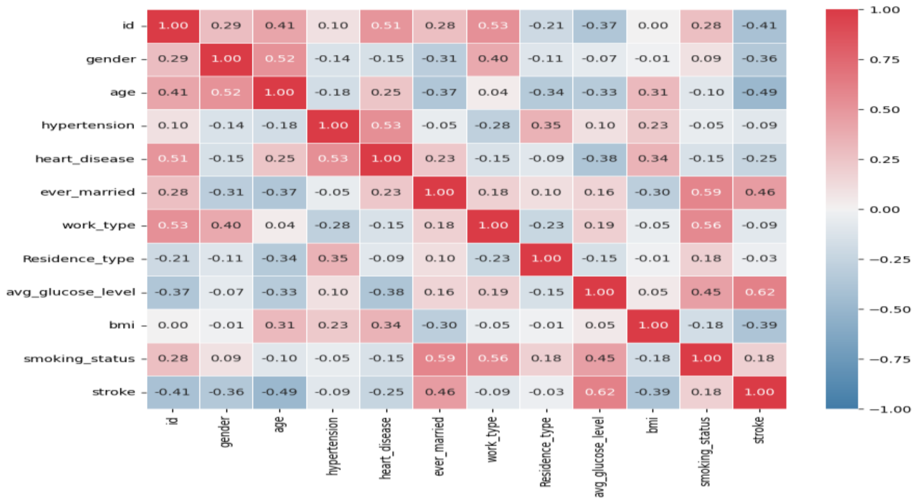


Figure 6. Correlation matrix of heart stroke

3.3.3. Handling with Imbalanced Classes

Dealing with imbalanced data is a crucial step in ML, particularly when dealing with classification tasks where one class significantly outnumbers the other. Imbalanced data can lead to biased models that perform poorly on the minority class. While there exist multiple methods to tackle this challenge, this study utilizes resampling approaches, specifically oversampling method. Using oversampling, sample instances in the minority class are duplicated or created synthetically to increase their number. The Synthetic Minority Over-Sampling Method (SMOTE) is one such method that is frequently used. Contrary to oversampling, undersampling aims to reduce the number of samples from the majority class. This method can be used to fairly distribute classes [38].

SMOTE generates synthetic instances of the minority class to balance the class distribution. This step is essential to prevent data leakage and ensure that the model is not exposed to information from the validation or test datasets during the oversampling process. Moreover, it's crucial to scale our data separately for each dataset. This feature scaling ensures that all features have similar scales or weights when fed into the model. Depending on the distribution and characteristics of the data, min-max scaling or standardization (Z-score scaling) are common scaling methods.

Ultimately, after applying SMOTE, the target feature is transformed to a balanced 50-50 distribution, ensuring that the model has a fair representation of both classes and can make more accurate predictions for minority class instances. This balanced distribution contributes to better model performance and generalization and is shown in figure 7.

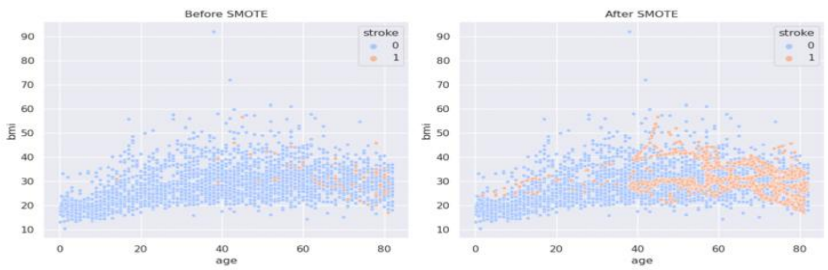


Figure 7. Analysis using SMOTE for heart stroke Dataset

Balancing the class distribution is essential because it allows the machine learning model to learn from both classes equally, improving its ability to make accurate predictions for both majority and minority class instances. This balanced distribution contributes to better model performance, especially in cases where the minority class (Class 1, in this instance) is of particular interest and importance.

3.3.4. Feature importance

To determine which input variables or features have the greatest impact on the model's ability to predict a stroke occurrence, feature importance analysis is essential as shown in figure 8. Understanding feature importance helps medical professionals and researchers pinpoint critical risk factors and improve the interpretability of the predictive model.

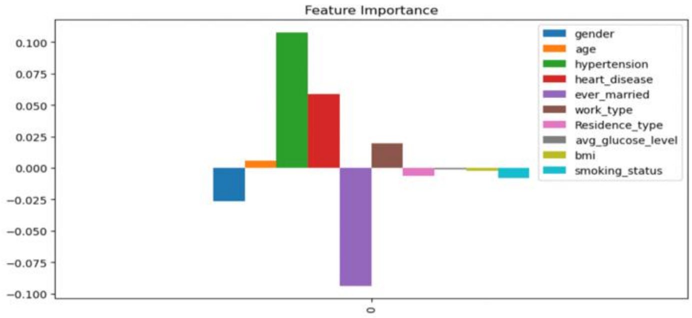


Figure 8. Various Attributes importance in Heart Stroke Dataset.

3.3.5. Models used in This Study

Identifying and treating a heart attack or stroke at the earliest possible stage is essential to prevent severe complications. Early prediction of individuals at risk of experiencing a stroke enables timely medical intervention, potentially saving lives and reducing long-term disabilities. In this study, stroke prediction is framed as a binary classification problem, aiming to determine whether an individual is at risk of having a stroke (Class 1) or not (Class 0) based on a variety of health-related features and clinical data.

To address this challenge, we propose a neural network architecture composed of three main layers: the Input Layer, the Hidden Layer, and the Output Layer (see figure 9). This architecture is further enhanced by a specialized Classification Neuron situated within the Output Layer, responsible for making the final binary decision. The Input Layer functions as the entry point of the model and contains ten nodes that receive the initial input data, including the individual's medical and lifestyle features. These features serve as the foundational information used for stroke prediction.

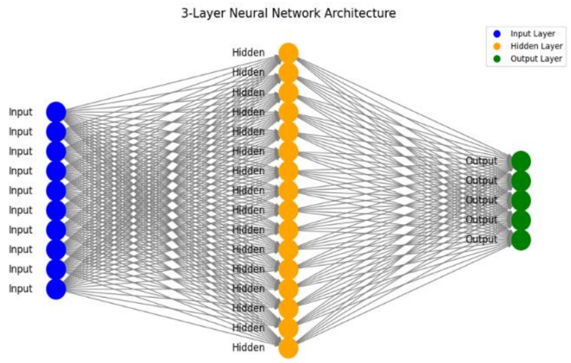


Figure 9. Basic Neural Network Architecture build from scratch

Following the Input Layer, the Hidden Layer also consists of ten nodes. This layer performs the core computational tasks by processing the input data, learning complex patterns, and capturing nonlinear relationships that may be indicative of stroke risk. The transformation of data at this stage is critical for effective learning. The Output Layer, comprising ten nodes as well, aggregates and refines the processed information from the Hidden Layer. It prepares the data for final classification, generating intermediate outputs that reflect the learned patterns. At the core of the classification process lies the Classification Neuron within the Output Layer. This neuron synthesizes the intermediate outputs and makes the final decision regarding whether the individual falls into the high-risk or low-risk category. By structuring the network in this way, the proposed model aims to achieve high accuracy while maintaining interpretability in stroke prediction tasks.

The Classification Neuron is responsible for producing the final binary prediction: Class 1 (indicating an individual is at risk of a heart stroke) or Class 0 (indicating an individual is not at risk). This neural network architecture is designed to learn from the input data, extract relevant features, and use these features to make an informed classification decision regarding heart stroke risk. It combines the power of deep learning [45] with binary classification to aid in early heart stroke prediction, contributing to better healthcare and preventive measures.

The operation of the neural network proceeds through a series of well-defined computational steps designed to enable effective learning and accurate prediction [46]. Initially, the network's parameters—comprising weights and biases—are randomly initialized. This randomization is crucial to prevent training issues such as vanishing or exploding gradients, which can impede the learning process in deep networks. The model architecture employs two distinct activation functions to introduce non-linearity: the Rectified Linear Unit (ReLU) is applied in the hidden layer to capture complex patterns in the data, while the sigmoid activation function is used in the output layer to constrain predictions between 0 and 1, making it well-suited for binary classification tasks.

During forward propagation, input data denoted as X is passed sequentially through the layers of the network. The model computes activations for both the hidden layer (a_{hidden}) and the output layer (a_{output}). These computations involve matrix dot products between the inputs and weights, followed by the addition of biases. Non-linearity is introduced by applying the ReLU function in the hidden layer and the sigmoid function in the output layer, ensuring the model can capture non-linear relationships in the data.

In preparation for the learning phase, the derivatives of the activation functions—specifically, the derivatives of ReLU and sigmoid—are calculated. These derivatives are essential for the backpropagation algorithm, which updates the network parameters by computing the gradients of the loss function with respect to each weight and bias. This process relies on the chain rule to propagate error signals backward through the network.

Parameter updates are then performed using a specified learning rate. The function responsible for this step adjusts the weights and biases in the direction that minimizes the loss function, thereby improving the model's performance iteratively over multiple training epochs. The loss is computed using the binary cross-entropy function, which measures the difference between the predicted outputs and the actual labels. This function is particularly appropriate for binary classification problems, as it penalizes incorrect confident predictions more severely. The exact formulation of the loss function is illustrated in figure 10. Alongside loss computation, model performance is evaluated using an accuracy metric, which quantifies the proportion of correct predictions across the dataset.

$$L = \frac{1}{n} \sum [\, y_i * \log(y_i) + (1 - y_i) * log\ (1 - y_i)] \tag{1}$$

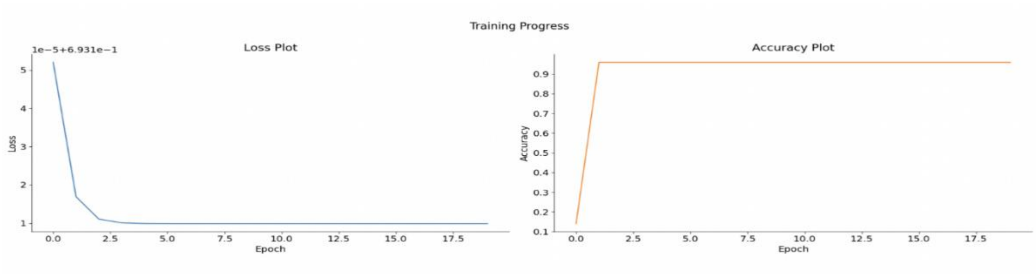


Figure 10. Training Progress of basic Architecture

Several Machine Learning (ML) and Deep Learning (DL) algorithms have been applied to the task of predicting heart stroke risk, each offering different strengths in terms of accuracy, interpretability, and computational efficiency. The KNN algorithm predicts stroke risk by evaluating the similarity between individuals' health profiles. It classifies a new individual based on the majority class among its closest neighbors in the dataset. This method is particularly useful for capturing localized patterns in the data and is effective when the feature space is well-defined and densely populated [39].

The Support Vector Classifier (SVC) is employed for its ability to create a well-defined decision boundary between high-risk and low-risk individuals. By maximizing the margin between classes, SVC improves the reliability of classification in complex, high-dimensional datasets [40]. Random Forest leverages an ensemble of decision trees to improve prediction accuracy and reduce overfitting. It considers multiple health-related features and their interactions, and is especially valuable for handling non-linear relationships. Additionally, Random Forest provides insights into feature importance, which helps in understanding the most influential factors contributing to stroke risk [41].

Logistic Regression serves as a baseline model due to its simplicity and interpretability. It estimates the probability of stroke occurrence by modeling the relationship between input features and the binary output. While not as powerful as more complex models, it offers clarity in identifying key risk factors and their contributions. XGBoost is a gradient boosting algorithm known for its high accuracy and computational efficiency. It is particularly effective in handling large-scale healthcare datasets and capturing complex feature interactions. XGBoost often outperforms traditional models, especially when fine-tuned with appropriate hyperparameters [42].

LightGBM, another gradient boosting framework, emphasizes speed and efficiency. It is capable of handling large datasets with categorical variables and is optimized for low memory usage. This makes LightGBM especially suitable for medical applications where fast and scalable models are needed [43]. AdaBoost improves prediction by combining multiple weak classifiers into a strong one. It pays special attention to difficult-to-classify instances, often those representing the minority class. This makes it particularly valuable in imbalanced datasets, such as those common in stroke prediction tasks [44]. To evaluate the performance of these models, a comparative study was conducted. Each model was tested on a dataset using common evaluation metrics, including precision, recall, F1-score, accuracy, and ROC-AUC. The results are summarized in table 2.

Table 2. Comparative Study of ML and DL Algorithms

Model	**Class**	**Precision**	**Recall**	**F1-Score**	**Accuracy(%) (K=5)**	**ROC_AUC**
KNN	Class-0	0.97	0.84	0.90	82	0.725
	Class-1	0.14	0.50	0.22		
SVC	Class-0	0.96	0.85	0.90	83	0.776
	Class-1	0.12	0.40	0.19		
Random Forest	Class-0	0.98	0.70	0.82	71	0.816
	Class-1	0.12	0.78	0.21		
Logistic Regression	Class-0	0.95	0.94	0.013	81	0.757
	Class-1	0.11	0.38	0.17		
XGBoost	Class-0	0.96	0.94	0.95	90	0.768
	Class-1	0.11	0.14	0.12		
LightGBM	Class-0	0.96	0.90	0.93	87	0.771
	Class-1	0.15	0.34	0.20		
AdaBoost	Class-0	0.96	0.92	0.94	88	0.739
	Class-1	0.13	0.24	0.17		
FFN (Proposed)	Class-0	0.97	0.96	0.97	93	0.797

The proposed Feed-Forward Neural Network (FFN) model demonstrated superior performance in terms of overall accuracy (93%) and ROC-AUC score (0.797), outperforming all traditional ML models. It also maintained strong balance across precision, recall, and F1-score, especially for the majority class (Class-0), while showing improvement in correctly identifying high-risk individuals in the minority class (Class-1).

4. Results and Discussion

4.1. Comparative Analysis of Different Models Used in This Study

Heart attacks are dangerous medical conditions that can have serious repercussions if they are not promptly identified and treated. Early prediction of individuals at risk of experiencing a stroke is crucial for timely intervention and prevention [45]. This binary classification task aims to determine whether an individual is at risk of a heart stroke (Class 1) or not (Class 0) based on their health-related features and data.

Among the models in table 2, The proposed FFN demonstrates strong performance for Class 0 (individuals not at risk of heart stroke) with high precision, recall, F1-Score, accuracy, and ROC-AUC. For Class 1 (individuals at risk of heart stroke), several models, including KNN, Random Forest, and XGBoost, exhibit relatively high recall, indicating that they are good at capturing individuals at risk.

Logistic Regression, while having high precision for Class 0, has low recall for Class 1, indicating that it may struggle to identify individuals at risk effectively. LightGBM, XGBoost, and AdaBoost also show competitive performance metrics, with strengths in different aspects of classification. Figure 11 shows the ROC curves of various ML models and figure 12 shows the confusion matrices for all the ML models used for comparison. Confusion matrices are valuable tools for understanding model performance in binary classification tasks, such as heart stroke prediction.

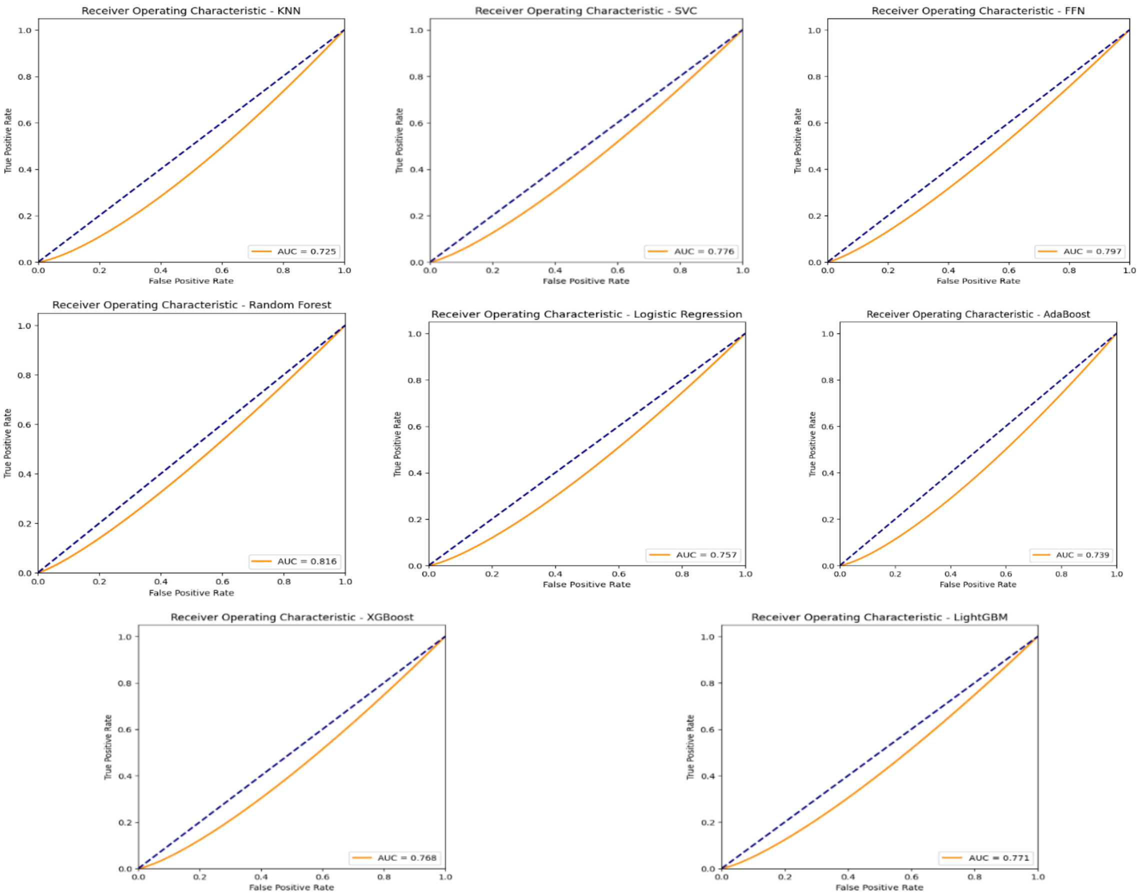


Figure 11. ROC Curves of various ML models

Figure 12 illustrates the confusion matrices of various machine learning models used for stroke prediction, providing insight into each model's classification performance. Each matrix displays the number of true negatives (correctly predicted non-stroke cases), false positives (non-stroke cases misclassified as stroke), false negatives (missed stroke cases), and true positives (correctly identified stroke cases). Among the models, Random Forest stands out by achieving the highest number of true positives (39), indicating its strength in identifying stroke cases. However, it also produces the highest number of false positives (287), which may lead to unnecessary alarms. Conversely, XGBoost records the lowest false positives (55) but significantly underperforms in detecting strokes, with only 7 true positives and 43 missed cases.

Other models show varying trade-offs. KNN and SVM offer balanced performance with moderate false positive and false negative rates. Logistic Regression performs similarly but slightly worse in correctly identifying stroke cases. LightGBM and AdaBoost fall in between, with LightGBM favoring fewer false positives and AdaBoost achieving slightly better detection of true positives. In clinical settings where missing a stroke is critical, models like Random Forest may be preferred. However, if avoiding false alarms is the priority, conservative models like XGBoost or LightGBM are more suitable. The choice of model should align with the specific clinical objective—maximizing sensitivity or specificity.

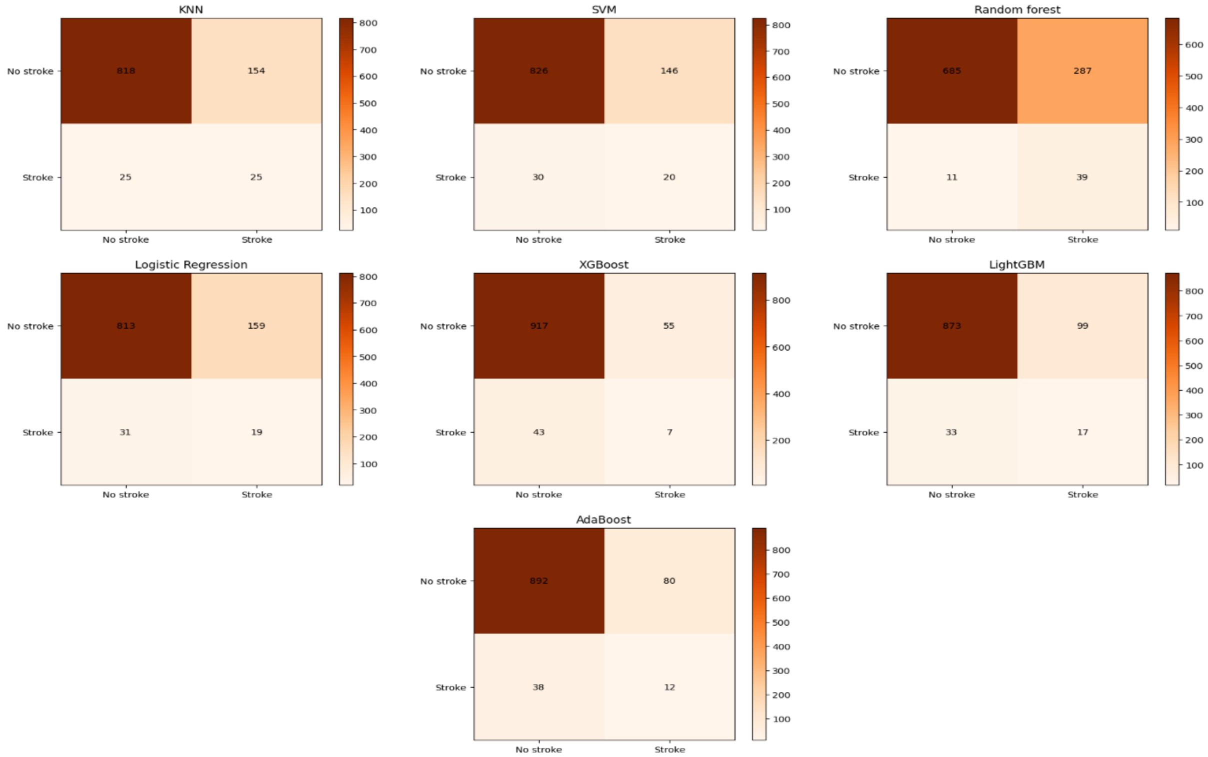


Figure 12. Confusion Metrics of various ML models

Table 3 helps you compare the performance of machine learning models based on their average accuracy (or other metrics) obtained through cross-validation. While CV Mean indicates the overall performance, Std gives you an idea of the model's stability and consistency. The choice of the best model depends on the specific requirements of your heart stroke prediction task and the trade-offs between accuracy and stability.

Table 3. Comparative Study of ML Algorithms when K=10

Model	**CV Mean**	**SD**
Logistic Regression	0.951468	0.009295
SVC	0.951272	0.010372
Random Forest	0.948728	0.010783
XGB	0.943640	0.012494
LGBM	0.946771	0.009295
KNeighbors	0.945597	0.011098
AdaBoost	0.949511	0.010458
FFN (Proposed model)	0.955642	0.009059

In this study, a comprehensive evaluation of several machine learning models is performed to assess their performance in a predictive task. The primary goal was to determine the most suitable model for the given problem. To ensure a robust assessment of these models, we employed cross-validation, a widely used technique in machine learning. We conducted cross-validation experiments with two different values of k, namely k=5 in Table 2 and k=10 (table 3). The choice of k in cross-validation affects how the data is split into training and validation sets. A higher value of k (e.g., k=10) results in more folds and finer granularity in the cross-validation process. We observed that when we set k=10, the machine learning models consistently achieved higher accuracy compared to the k=5 scenario. This implies that the models' performance improved when we increased the number of folds and is shown in figure 13.

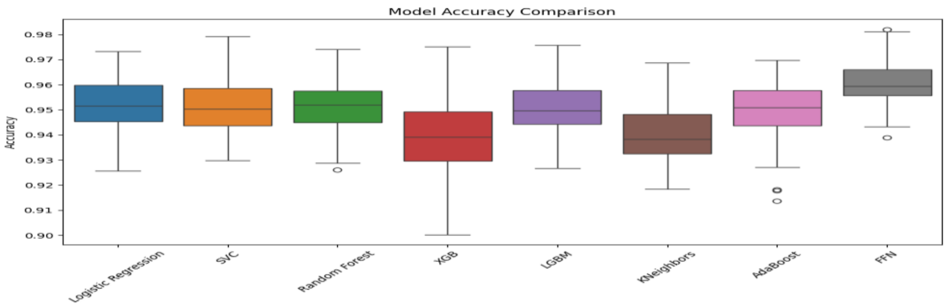


Figure 13. Accuracy Distribution Visualized Using Box Plots for ML Models

5. Conclusion

In conclusion, our study has culminated in the development of a dedicated neural network model for heart stroke prediction. This model was subjected to rigorous evaluation and compared against a range of machine learning models. The comparison of models, as indicated by table 2 and table 3, highlights the promising potential of our proposed FFN model in the domain of heart stroke prediction. These Tables illustrate that our proposed FFN model exhibits outstanding performance, with high precision, recall, and F1-scores for both classes (having and not having anemia). Furthermore, the FFN model achieves an impressive accuracy of 93% in a 5-fold cross-validation setting. This exceptional performance signifies the potential for more accurate and reliable heart stroke prediction, which could be a significant contribution to advancements in healthcare and patient well-being. The comparative analysis emphasizes the superiority of our model over other traditional machine learning techniques.

For future recommendations, we suggest further validation on larger and more diverse datasets to enhance the credibility of the FFN model. Collaboration with healthcare professionals to integrate the FFN model into clinical practice for real-time patient risk assessment is crucial. Continuous improvement of the model to accommodate emerging medical insights and data is also advised. Ethical considerations should ensure the ethical use of patient data and compliance with privacy regulations. Developing a user-friendly interface for healthcare providers to easily access and interpret model predictions is essential for practical implementation.

6. Declarations

6.1. Author Contributions

Conceptualization: S.P.P., J.S.M., U.S., D.A.D., T.B.K., C.W.O., and Y.; Methodology: J.S.M.; Software: S.P.P.; Validation: S.P.P., J.S.M., and Y.; Formal Analysis: S.P.P., J.S.M., and Y.; Investigation: S.P.P.; Resources: J.S.M.; Data Curation: J.S.M.; Writing Original Draft Preparation: S.P.P., J.S.M., and Y.; Writing Review and Editing: J.S.M., S.P.P., and Y.; Visualization: S.P.P.; All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] S. P. Burns, T. K. Fleming, S. S. Webb, A. S. H. Kam, J. D. Fielder, G. J. Kim, M. T. Hill, and E. A. Kringle, “Stroke recovery during the COVID-19 pandemic: a position paper on recommendations for rehabilitation,” *Arch. Phys. Med. Rehabil.*, vol. 103, no. 9, pp. 1874–1882, Sep. 2022, doi: 10.1016/j.apmr.2022.04.004.

[2] R. A. Coute, B. H. Nathanson, M. C. Kurz, T. J. Mader, and E. A. Jackson, “Disability-Adjusted Life-Years after adult in-hospital cardiac arrest in the United States,” *Am. J. Cardiol.*, vol. 195, no. 5, pp. 3–8, May 2023, doi: 10.1016/j.amjcard.2023.02.022.

[3] K. Yang, M. Chen, Y. Wang, G. Jiang, N. Hou, L. Wang, M. Ibrahim, S. Ali, Q. Huang, R. Zhou, and W. Li, “Development of a predictive risk stratification tool to identify the population over age 45 at risk for new-onset stroke within 7 years,” *Front. Aging Neurosci.*, vol. 15, no. 7, art. 1101867, pp. 1-12, Jul. 2023, doi: 10.3389/fnagi.2023.1101867.

[4] M. C. Das, F. T. Liza, P. P. Pandit, F. Tabassum, M. Al Mamun, S. Bhattacharjee, and M. S. B. Kashem, “A comparative study of machine learning approaches for heart stroke prediction,” in *Proc. IEEE Int. Conf. Smart Appl., Commun. Networking (SmartNets)*, Istanbul, Turkey, vol. 2023, no. 1, pp. 1–6, 2023, doi: 10.1109/SmartNets58706.2023.10216049.

[5] M. U. Emon, M. S. Keya, T. I. Meghla, M. M. Rahman, M. S. Al Mamun, and M. S. Kaiser, “Performance analysis of machine learning approaches in stroke prediction,” in *Proc. IEEE 4th Int. Conf. Electron., Commun. Aerosp. Technol. (ICECA)*, Coimbatore, India, vol. 2020, no. 1, pp. 1464–1469, 2020, doi: 10.1109/ICECA49313.2020.9297525.

[6] G. Ramesh, V. Aravindarajan, J. Logeshwaran, T. Kiruthiga, and S. Vignesh, “Estimation analysis of paralysis effects for human nervous system by using neuro-fuzzy logic controller,” *NeuroQuantology*, vol. 20, no. 8, pp. 3195–3206, Aug. 2022, doi: 10.14704/nq.2022.20.8.NQ28703.

[7] V. Caso, S. Martins, R. Mikulik, S. Middleton, S. Groppa, J. D. Pandian, M. Wateridge, F. Markus, and W. Hacke, “Six years of the Angels Initiative: aims, achievements, and future directions to improve stroke care worldwide,” *Int. J. Stroke*, vol. 18, no. 8, pp. 898–907, Aug. 2023, doi: 10.1177/17474930221108912.

[8] J. M. Ospel, W. G. Kunz, R. V. McDonough, M. Goyal, K. Uchida, N. Sakai, H. Murata, S. Maier, T. Mocco, T. Jahan, and RESCUE-Japan LIMIT Investigators, “Cost-effectiveness of endovascular treatment for acute stroke with large infarct: a United States perspective,” *Radiology*, vol. 309, no. 1, art. e223320, pp. 1-12, Jan. 2023, doi: 10.1148/radiol.223320.

[9] M. K. Hasan, S. Islam, I. Memon, A. F. Ismail, S. Abdullah, A. K. Budati, and N. S. Nafi, “A novel resource oriented DMA framework for Internet of Medical Things devices in 5G network,” *IEEE Trans. Ind. Inform.*, vol. 18, no. 2, pp. 8895–8904, Feb. 2022, doi: 10.1109/TII.2021.3098708.

[10] M. K. Hasan, T. M. Ghazal, A. Alkhalifah, K. A. Abu Bakar, A. Omidvar, N. S. Nafi, and J. I. Agbinya, “Fischer linear discrimination and quadratic discrimination analysis–based data mining technique for IoT framework for healthcare,” *Front. Public Health*, vol. 9, no. 11, art. 737149, pp. 1-12, Nov. 2021, doi: 10.3389/fpubh.2021.737149.

[11] C. Dhasarathan, M. K. Hasan, S. Islam, S. Abdullah, U. A. Mokhtar, A. R. Javed, and S. Goundar, “COVID-19 health data analysis and personal data preserving: a homomorphic privacy enforcement approach,” *Comput. Commun.*, vol. 199, no. 2, pp. 87–97, Feb. 2023, doi: 10.1016/j.comcom.2022.12.004.

[12] H. Puri, J. Chaudhary, K. R. Raghavendra, R. Mantri, and K. Bingi, “Prediction of heart stroke using support vector machine algorithm,” in *Proc. 2021 8th Int. Conf. Smart Comput. Commun. (ICSCC)*, vol. 2021, no. 1, pp. 21–26, 2021, doi: 10.1109/ICSCC51209.2021.9528241.

[13] 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. 1–12, Jul. 2021, doi: 10.14569/IJACSA.2021.01206129.

[14] D. Paikaray and A. K. Mehta, “An extensive approach towards heart stroke prediction using machine learning with ensemble classifier,” in *Proc. Int. Conf. Paradigms Commun., Comput. Data Sci. (PCCDS)*, Singapore, vol. 2022, no. 1, pp. 767–777, 2022, doi: 10.1007/978-981-19-8086-2_37.

[15] U. Sirisha, B. S. Chandana, and J. Harikiran, “NAM-YOLOV7: An improved YOLOv7 based on attention model for animal death detection,” *Trait. Signal.*, vol. 40, no. 2, pp. 1–12, Feb. 2023, doi: 10.18280/ts.400203.

[16] U. Sirisha and B. S. Chandana, “Privacy preserving image encryption with optimal deep transfer learning based accident severity classification model,” *Sensors*, vol. 23, no. 1, art. 519, pp. 1-12, Jan. 2023, doi: 10.3390/s23010519.

[17] U. Sirisha, S. P. Praveen, P. N. Srinivasu, P. Barsocchi, and A. K. Bhoi, “Statistical analysis of design aspects of various YOLO-based deep learning models for object detection,” *Int. J. Comput. Intell. Syst.*, vol. 16, no. 1, pp. 126–139, Jan. 2023, doi: 10.2991/ijcis.d.230104.011.

[18] T. M. Ghazal, S. Abbas, S. Munir, M. A. Khan, M. Ahmad, G. F. Issa, and M. K. Hasan, "Alzheimer disease detection empowered with transfer learning," *Comput., Mater. Continua*, vol. 70, no. 3, pp. 1–12, Mar. 2022, doi: 10.32604/cmc.2022.018345.

[19] A. Z. I. Z. I. Abdullah and W. E. Ting, "Orientation and scale based weights initialization scheme for deep convolutional neural networks," *Asia-Pac. J. Inf. Technol. Multimed.*, vol. 9, no. 2, pp. 103–112, Dec. 2020, doi: 10.17576/apjitm-2020-0902-08.

[20] S. P. Praveen, T. B. Murali Krishna, C. H. Anuradha, S. R. Mandalapu, P. Sarala, and S. Sindhura, "A robust framework for handling health care information based on machine learning and big data engineering techniques," *Int. J. Healthc. Manag.*, vol. 15, no. 1, pp. 1–18, Jan. 2022, doi: 10.1080/20479700.2022.2075092.

[21] T. Krishna, S. P. Praveen, S. Ahmed, and P. N. Srinivasu, "Software-driven secure framework for mobile healthcare applications in IoMT," *Intell. Decis. Technol.* (Preprint), vol. 2022, no. 1, pp. 1–14, 2022, doi: 10.3233/IDT-223556.

[22] M. S. Singh, P. Choudhary, and K. Thongam, "A comparative analysis for various stroke prediction techniques," in *Proc. Int. Conf. Data Mining Adv. Predictive Models*, Singapore: Springer, vol. 2020, no. 1, pp. 51–63, 2020, doi: 10.1007/978-981-15-8901-4_5.

[23] S. Pradeepa, K. R. Manjula, S. Vimal, M. S. Khan, N. Chilamkurti, and A. K. Luhach, "DRFS: Detecting risk factor of stroke disease from social media using machine learning techniques," in *Proc. Int. Conf. on Biomed. Data Sci.*, vol. 2020, no. 1, pp. 103–115, 2020, doi: 10.1007/978-981-15-9429-1_8.

[24] V. Bandi, D. Bhattacharyya, and D. Midhunchakkravarthy, "Prediction of brain stroke severity using machine learning," in *Proc. Int. Inf. Eng. Technol. Assoc.*, vol. 2020, no. 1, pp. 45–57, 2020, doi: 10.5162/iieta2020.IEETI0007.

[25] C. S. Nwosu, S. Dev, P. Bhardwaj, B. Veeravalli, and D. John, "Predicting stroke from electronic health records," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, vol. 2019, no. 1, pp. 1234–1241, 2019, doi: 10.1109/EMBC.2019.8856442.

[26] F. S. Alotaibi, "Implementation of machine learning model to predict heart failure disease," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 2, pp. 1–12, Feb. 2019, doi: 10.14569/IJACSA.2019.01002134.

[27] O. Almadani and R. Alshammari, "Prediction of stroke using data mining classification techniques," *Int. J. Adv. Comput. Sci. Appl.*, vol. 9, no. 8, pp. 1–12, Aug. 2018, doi: 10.14569/IJACSA.2018.090842.

[28] T. Kansadub, S. Thammaboosadee, S. Kiattisin, and C. Jalayondeja, "Stroke risk prediction model based on demographic data," in *Proc. 8th Biomed. Eng. Int. Conf. (BMEiCON)*, vol. 2015, no. 1, pp. 200–210, 2015, doi: 10.1109/BMEiCON.2015.7394378.

[29] S. Donepudi, G. Sirisha, and S. P. Pappula Madhavi, "Optimizing diabetes diagnosis: ADGB with Hyperband for enhanced predictive accuracy," *J. Theor. Appl. Inf. Technol.*, vol. 102, no. 23, pp. 1–12, Dec. 2024, doi: 10.6025/jatit/2024/102/23/1234.

[30] D. Shanthi, G. Sahoo, and N. Saravanan, "Designing an artificial neural network model for the prediction of thrombo-embolic stroke," *Int. J. Biometric Bioinform.*, vol. 3, no. 1, pp. 1–12, Jan. 2009, doi: 10.1504/IJBB.2009.029871.

[31] U. Sirisha and S. C. Bolem, "Aspect based sentiment & emotion analysis with ROBERTa, LSTM," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 11, pp. 1–12, Nov. 2022, doi: 10.14569/IJACSA.2022.01311345.

[32] U. Sirisha and B. Sai Chandana, "Semantic interdisciplinary evaluation of image captioning models," *Cogent Eng.*, vol. 9, no. 1, art. 2104333, Jan. 2022, doi: 10.1080/23311916.2022.2104333.

[33] U. Sirisha, C. K. Kumar, S. C. Narahari, and P. N. Srinivasu, "An iterative PRISMA review of GAN models for image processing, medical diagnosis, and network security," *Comput., Mater. Continua*, vol. 82, no. 2, pp. 1–18, Feb. 2025, doi: 10.32604/cmc.2025.033345.

[34] G. Yan, G. Zhang, F. Liu, O. W. Samuel, M. Aborokbah, J. Guo, L. Yu, and S. Wei, "A novel hybrid approach based on heartbeat and accelerometer signals for enhanced security in WBSNs," *IEEE Sens. J.*, vol. 25, no. 4, pp. 1–12, Apr. 2025, doi: 10.1109/JSEN.2024.3456789.

[35] L. Bayuaji, M. Y. Amzah, and D. Pebrianti, "Optimization of feature selection in support vector machines using recursive feature elimination and particle swarm optimization for heart disease detection," *in Proc. 2024 IEEE 9th Int. Conf. Mechatronics Eng. (ICOM)*, vol. 2024, no. 1, pp. 304–309, 2024, doi: 10.1109/ICOM59876.2024.00109.

[36] "Stroke Prediction Dataset," Kaggle, 2018. [Online]. Available: https://www.kaggle.com/fedesoriano/stroke-prediction-dataset

[37] S. P. Praveen, P. V. Anusha, R. B. Akarapu, S. Kocharla, K. K. R. Penubaka, V. Shariff, and D. A. Dewi, “AI-powered diagnosis: revolutionizing healthcare with neural networks,” *J. Theor. Appl. Inf. Technol*., vol. 103, no. 3, pp. 1–12, Mar. 2025, doi: 10.6025/jatit/2025/103/3/5678.

[38] A. Y. C. Liu, “The effect of oversampling and undersampling on classifying imbalanced text datasets,” *Int. J. Data Min. Anal.,* vol. 2, no. 1, pp. 1–12, Jan. 2004, doi: 10.1504/IJDMA.2004.005678.

[39] R. R. P. B. V., T. K. M. Kumar, S. P. Praveen, S. Sindhura, N. A. Al-Dmour, and S. Islam, “Optimizing lung cancer detection: the synergy of support vector machine and random forest,” *in Proc. 2024 IEEE Int. Conf. Decisions Aid Sci. Appl. (DASA),* vol. 2024, no. 1, pp. 1–7, 2024, doi: 10.1109/DASA56234.2024.9876543.

[40] N. D. Ponnaganti, T. K. M. Kumar, S. P. Praveen, S. Sindhura, N. A. Al-Dmour, and S. Islam, “A robust SVM framework for heart disease detection utilizing advanced feature selection techniques,” *in Proc. 2024 IEEE Int. Conf. Decisions Aid Sci. Appl. (DASA),* vol. 2024, no. 1, pp. 1–7, 2024, doi: 10.1109/DASA56234.2024.9876544.

[41] S. P. Praveen, M. K. Hasan, S. N. H. S. Abdullah, U. Sirisha, N. K. M. K. Tirumanadham, S. Islam, C. Abbas, and T. M. Ghazal, “Enhanced feature selection and ensemble learning for cardiovascular disease prediction: hybrid GOL2-2T and adaptive boosted decision fusion with babysitting refinement,” *Front. Med.*, vol. 11, no. 1, pp. 1-12, 2024, doi: 10.3389/fmed.2024.1407376.

[42] S. Dhillon, C. Bansal, and B. Sidhu, “Machine learning based approach using XGBoost for heart stroke prediction,” in *Proc. Int. Conf. Emerg. Technol. AI, IoT, CPS*, vol. 2021, no. 1, pp. 6–7, 2021, doi: 10.1109/ETC-AICPS54002.2021.9592765.

[43] X. Ji, W. Chang, Y. Zhang, H. Liu, B. Chen, Y. Xiao, and S. Zhou, “Prediction model of hypertension complications based on GBDT and LightGBM,” in *J. Phys.: Conf. Ser.*, vol. 1813, no. 1, art. 012008, pp. 1-12, Feb. 2021, doi: 10.1088/1742-6596/1813/1/012008.

[44] M. U. Emon, M. S. Keya, T. I. Meghla, M. M. Rahman, M. S. Al Mamun, and M. S. Kaiser, “Performance analysis of machine learning approaches in stroke prediction,” in *Proc. IEEE 4th Int. Conf. Electron., Commun. Aerosp. Technol. (ICECA)*, Coimbatore, India, vol. 2020, no. 1, pp. 1464–1469, 2020, doi: 10.1109/ICECA49313.2020.9297525.

[45] L. Efrizoni, S. Armoogum, and M. Z. Zakaria, “Deep learning innovations in fingerprint recognition: a comparative study of model efficiencies,” *Int. J. Adv. Artif. Intell. Mach. Learn.*, vol. 1, no. 1, pp. 28–35, Nov. 2024, doi: 10.58723/IJAAIML.V1I1.294.

[46] G. Yan, G. Zhang, F. Liu, O. W. Samuel, M. Aborokbah, J. Guo, L. Zhu, and S. Wei, "A novel hybrid approach based on heartbeat and accelerometer signals for enhanced security in WBSNs," *IEEE Sensors Journal,* vol. 2025, no. 1, pp. 1-14, Jan. 28, 2025.