COMPUTING AND ALGORITHM INSIGHT
Volume 3, Issue 2, December 2025 | pp. 99-116
https://doi.org/10.63208/21018-117

OPEN ACCESS
Check for updates

AN INTELLIGENT STROKE DETECTION FRAMEWORK UTILIZING GENETIC ALGORITHM-BASED FEATURE SELECTION AND BIDIRECTIONAL LSTM ON CT IMAGERY

Sarah Mensah[1] ✉, Li Xiangyu[1], and Nguyen Hoang Long[2]

School of Life Sciences, Singapore University of Social Sciences (SUSS), Singapore[1]; Faculty of Biology and Biotechnology, University of Science - VNUHCM, Ho Chi Minh City, Vietnam[2]

ABSTRACT

Cerebrovascular pathologies, particularly strokes, are a leading cause of global mortality and morbidity, where timely diagnosis is critical for effective clinical intervention. This research introduces an advanced diagnostic framework for early stroke identification by integrating Computed Tomography (CT) imaging with a hybrid computational approach. The system utilizes a Genetic Algorithm (GA) to optimize feature selection from neural network outputs, identifying the most discriminative biomarkers within the image data. These refined features are subsequently classified using a Bidirectional Long Short-Term Memory (BiLSTM) network. Rigorous validation through cross-validation showed that the proposed model achieved a 96.5% accuracy, significantly outperforming traditional classifiers like Random Forests and Support Vector Machines. By utilizing metrics such as AUC-ROC and F1-score, the study demonstrates that this architecture serves as a robust decision-support tool, enabling medical practitioners to implement more precise and rapid stroke management strategies.

Keywords: *Stroke Detection, Genetic Algorithm, Bidirectional LSTM, CT Image Classification, Feature Selection.*

CORRESPONDING AUTHOR:
Sarah Mensah
School of Life Sciences, Singapore University of Social Sciences (SUSS), Singapore
contact: *sarahmensah@suss.edu.sg*

ARTICLE HISTORY
Received : June 18, 2025
Final Revised : November 10, 2025
Accepted : December 4, 2025
Published : December 30, 2025

© 2025 The Author(s) | Published by Integra Academic Press | This article is an open access article and distributed under the terms and conditions of the Creative Commons Attribution (CC BY-SA) license.

1. | INTRODUCTION

A stroke is precipitated by either an obstruction in the cranial blood supply or the rupture of a cerebral blood vessel. Globally, it remains the leading cause of mortality, accounting for more than 6.2 million annual deaths [1]. Beyond its lethality, survivors frequently experience significant disabilities that drastically diminish their quality of life. Implementing preventive strategies and ensuring rapid clinical intervention are essential for mitigating stroke severity and its long-term consequences. Consequently, the early identification of high-risk individuals is fundamental to effective stroke prevention [2]. Strokes are categorized into two primary types: ischemic and hemorrhagic. Ischemic strokes are the most prevalent, representing approximately 87% of all cases [3]. This condition occurs when blood vessels supplying the brain become obstructed or narrowed, thereby restricting or halting blood flow to specific regions. Such blockages are typically the result of blood clots, cholesterol buildup, or various embolic debris originating from other parts of the body [4]. Although hemorrhagic strokes occur less frequently, they are notably more severe and often prove fatal without immediate treatment. This type of stroke occurs when a cerebral vessel leaks or ruptures, releasing blood into the surrounding parenchyma, which induces swelling and exerts pressure on brain structures, ultimately damaging cells and impairing neurological function [5].

This research utilizes machine learning methodologies to establish a systematic framework for stroke identification. The primary objective is to deepen the understanding of stroke etiology and construct robust detection models that facilitate informed clinical decision-making regarding treatment and prevention. The study integrates the analysis of electronic health records, genetic markers, and lifestyle variables to build predictive architectures capable of identifying vulnerable individuals. By employing feature selection algorithms, the study aims to extract critical insights from complex datasets to enhance diagnostic precision. Furthermore, the resulting machine learning models are designed to function as sophisticated support tools for clinicians.

The overarching goal of this investigation is to assess the efficacy, interpretability, adaptability, and scalability of the developed algorithms. To determine comparative performance, the proposed models were benchmarked against established clinical risk assessment methods using multiple empirical datasets. The findings offer significant implications for clinical practice and the development of personalized interventions intended to reduce the burden of stroke on both patients and the healthcare infrastructure. Ultimately, this study applies advanced machine learning to forecast stroke frequency and mitigate occurrence through improved model accuracy and transparency. This approach has the potential to transform stroke prevention by facilitating precise, early-stage interventions. The primary contributions of this research are summarized as follows:

1) The development of a diagnostic architecture for stroke detection utilizing image-based datasets.
2) The implementation of a genetically optimized CNN to extract high-utility features from CT imaging.
3) A comparative analysis of LSTM, BiLSTM, and genetically optimized CNNs for stroke classification based on the extracted features.
4) A performance benchmark comparing the proposed GA_BiLSTM diagnostic system against other established machine learning (ML) and deep learning (DL) paradigms.

5) The provision of an early detection system designed to assist healthcare providers in making data-driven decisions for stroke management and prevention.

2. | RELATED WORKS

Machine learning algorithms have been instrumental in developing high-fidelity stroke risk detection models. These frameworks utilize longitudinal data from electronic health records to forecast stroke susceptibility, particularly in hypertensive cohorts [6]. Various architectures, such as Extreme Gradient Boosting (XGBoost), have been implemented to attain peak predictive performance [7, 8]. Furthermore, the deployment of ensemble models, which aggregate diverse computational approaches, has demonstrated significant potential in identifying high-risk individuals [9].

Regarding image preprocessing, the authors of [10] provide an extensive evaluation of skull stripping methodologies, contrasting contemporary deep learning-based techniques with classical approaches. This review highlights specific procedures that are highly compatible with routine clinical imaging workflows. In a different domain of predictive 101odelling, research in [11] introduces a temporal link prediction framework named TDRL, utilizing deep reinforcement learning to analyze evolving datasets. Their findings indicate that models trained on temporal data outperform those based on static time points, although the TDRL-CAN variant struggled with predicting disappearing edges in network topologies. This challenge of incomplete data is common in criminal network analysis (CAN), where most supervised techniques falter; however, the application of deep reinforcement learning (DRL) has proven superior to traditional supervised learning for predicting hidden network connections [12].

Specific to clinical neuroimaging, a study in [13] utilized a CNN hybrid structure derived from the National Institutes of Health to predict various stroke outcomes, such as the Modified Rankin Scale (mRS) score, achieving a 74% accuracy rate for 90-day prognosis. Another approach integrated wavelet entropy and spider network graphs with probabilistic neural networks to categorize MRI scans into normal, stroke-affected, or degenerative states [14]. By employing a 2D discrete wavelet transform for initial processing and spider web diagrams for feature extraction, they reported an exhaustive classification accuracy of 100%.

Analysis of electronic health records in [15] and [16] seeks to refine stroke management by identifying key risk factors—such as age, cardiac disease, glycemic levels, and hypertension—through principal component analysis. Their findings suggest that a perceptron neural network utilizing these four specific attributes yields the highest accuracy. They also mitigated dataset imbalances through sub-sampling techniques. Similarly, a deep learning framework discussed in [17] achieved a 96.21% average accuracy in detecting intracranial hemorrhage (ICH) subtypes. This system significantly lowers false positive rates and includes a quantitative scoring algorithm for measuring lesion volume and thickness, which is vital for surgical decision-making.

Hybrid feature selection strategies have also been explored, such as the use of discrete wavelet transforms (DWT), discrete cosine transforms (DCT), and Gray-Level Co-Occurrence Matrix (GLCM) features [18]. In this context, Random Forest classifiers attained an accuracy of 87.97% when combining DWT and GLCM features. A similar approach in [19] utilized SMOTE for oversampling and sequential forward selection to refine feature subsets, with Random Forest again leading at 87.22% accuracy. Furthermore, research in [20] demonstrated

that a CNN-RNN framework, supported by robust preprocessing, markedly improves ICH detection on CT scans, serving as an effective decision-support tool for radiologists.

Despite these advancements, the literature review indicates that many extant models either suffer from suboptimal accuracy or prohibitive computational demands. To address these limitations, we propose a diagnostic system that integrates a genetic algorithm for optimized feature selection with a Bidirectional Long Short-Term Memory (BiLSTM) network to analyze CT brain imagery. Our methodology is specifically designed to overcome the accuracy bottlenecks and high computational complexity inherent in previously established models.

3. | METHODOLOGY

This section elucidates the experimental methodology and the computational framework developed for this study. The conceptual architecture of the proposed computer-aided stroke diagnosis system is delineated in Figure 1. The process begins by loading the benchmark dataset into the system and partitioning it into training and validation subsets via the holdout cross-validation method. Specifically, 70% of the images across five distinct classes are randomly assigned to train the model, while the remaining 30% of the labeled data is reserved as the validation set to assess the system's performance.

To derive robust and non-invariant features from the RGB imaging data, five sophisticated convolutional neural network (CNN) architectures are utilized: AlexNet, NASNet-Large, VGG-19, Inception V3, and ShuffleNet. Each of these models processes the training and validation images to extract a comprehensive feature vector consisting of 1000 attributes from the final fully connected layer. Standard CNN architectures typically comprise two functional segments: feature extraction and classification; however, for our specific feature extraction pipeline, the models are utilized primarily to generate the high-dimensional feature vectors. Following extraction, a Genetic Algorithm (GA) employing tournament selection is implemented to rank the features based on their optimization potential. The feature vectors from both the training and validation sets are then meticulously analyzed, allowing for the systematic removal of non-optimal attributes and the retention of the most discriminative data points.

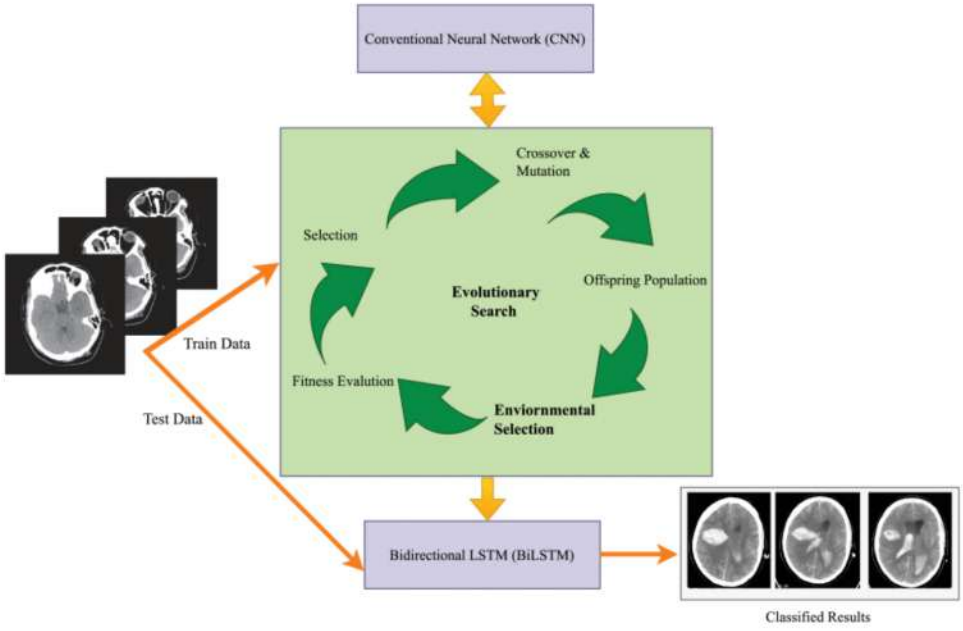


Figure 1. Framework of the proposed model.

2.1 CNN Architecture

Convolutional Neural Networks (CNNs) represent a specialized class of deep learning architectures primarily engineered for high-fidelity image recognition and pixel-level analysis [21, 22]. In the context of medical imaging, diverse CNN models—including AlexNet, Inception V3, NASNet-Large, ShuffleNet, and VGG-19—leveraging transfer learning techniques [23, 24] have demonstrated significant efficacy in classifying complex pathologies such as diabetic retinopathy. As illustrated in Figure 1, our framework integrates these pre-trained models with genetic algorithms to facilitate a systematic and optimized feature selection process [25].

2.1.1 AlexNet

Originally introduced by Alex Krizhevsky in 2012, AlexNet is a landmark CNN architecture characterized by approximately 650,000 neurons and 60 million trainable parameters. The model's internal composition is built upon a sequence of convolutional filters, Rectified Linear Unit (ReLU) activation functions, and fully connected layers, culminating in a softmax activation for classification. The efficiency and performance of AlexNet can be modulated by fine-tuning the depth of the layers and the density of the filters. Furthermore, specific architectural modifications, such as altering the connectivity between layers or adjusting input resolutions, can be implemented to vary the strength of feature representation. As fundamental components of neural networks, these layers rely on learned weights and biases to transform data. The structural design of AlexNet is presented in Figure 2, which provides a comprehensive overview of the data flow—from the initial input through weighted summation and activation to the final neuronal transmission and output.

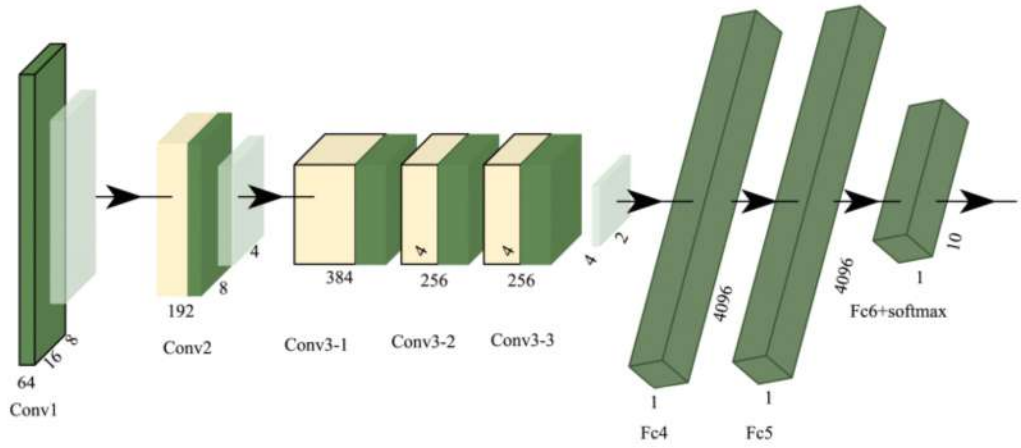


Figure 2. Framework of the proposed model.

2.1.2 Inception V3

In 2014, Google introduced a CNN architecture designed to estimate the local sparsity of a network by aggregating multiple specialized modules. By utilizing these sophisticated components, the architecture can effectively determine the optimal local sparsity within the network structure. In December 2015, Szegedy unveiled an enhanced iteration known as Inception V3, which demonstrated performance significantly superior to established benchmarks. This model exhibits versatile capabilities, including object identification, human action recognition, video segmentation, and the tracking of stationary and dynamic objects. Inception V3 employs convolutional factorizations and auxiliary classifiers to minimize the

grid size to its absolute limit. To mitigate representational bottlenecks before reaching maximum or average pooling layers, the network scales its filtering activations. Furthermore, the integration of auxiliary classifiers serves to bolster classification accuracy while providing deeper transparency into the internal network dynamics. The structural configuration of Inception V3 is depicted in Figure 3.

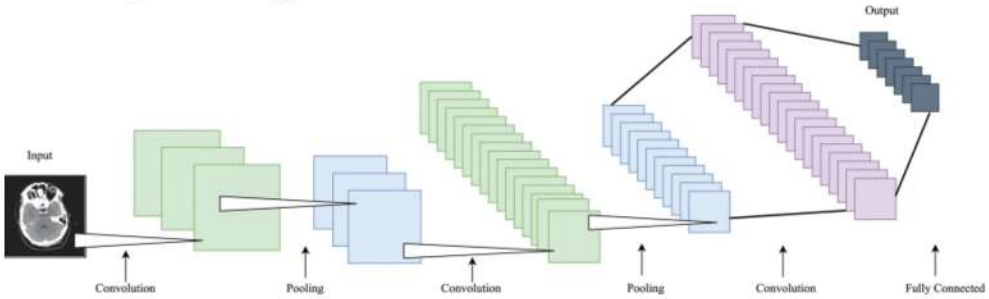


Figure 3. Pre-trained Inception V3 architecture.

2.1.3 VGG 19

Various architectures have been utilized to analyze the CNN-based ImageNet dataset, most notably the VGGNet series, developed by the Visual Geometry Group. Figure 4 illustrates the VGG-19 configuration, a streamlined yet highly versatile architecture designed to progressively reduce data volume through its hierarchical structure. This model has gained significant traction in both clinical research and commercial imaging sectors due to its distinctive arrangement of stacked convolutional layers followed by max-pooling stages, which effectively compress feature maps while extracting high-level abstractions.

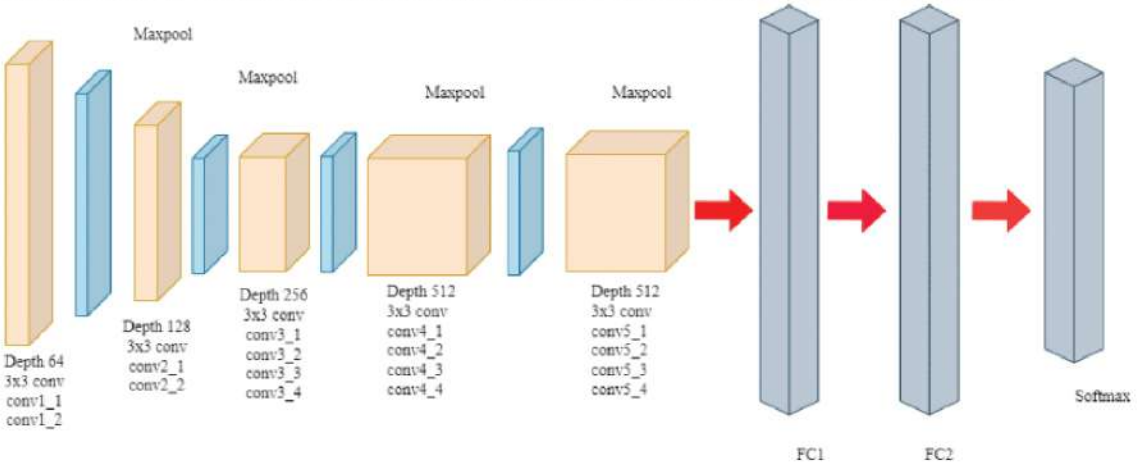


Figure 4. Pre-trained VGG 19 architecture.

2.1.4 NASNet-Large

Artificial neural networks (ANNs), while providing profound computational insights, are frequently influenced by human design biases, which can introduce structural limitations into their architectures [26]. To circumvent these constraints, NASNet-Large was developed as an automated architecture search algorithm by Google to identify the most efficient combination of features within a vast search space. As demonstrated in Figure 5, this model yields exceptional performance when deployed on large-scale datasets such as ImageNet, leveraging its reinforcement learning-based design to optimize network topology.

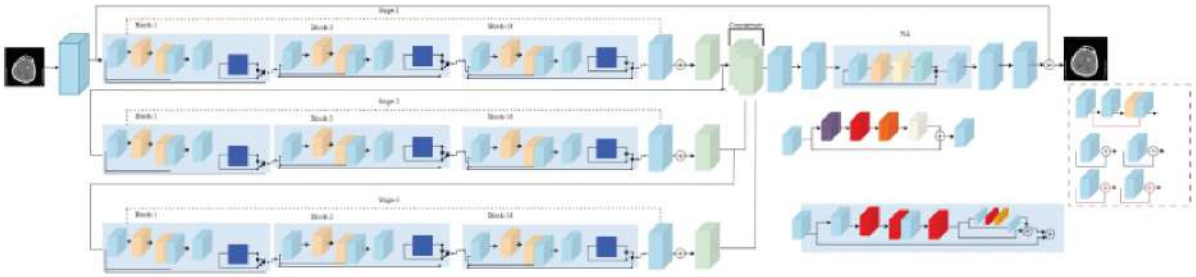


Figure 5. Pre-trained NASNet-Large architecture.

2.1.5 ShuffleNet

Introduced in 2017 by researchers at Megvii Inc., ShuffleNet was specifically engineered to facilitate high-performance deep learning on mobile and embedded devices with limited computational resources. The architectural details of ShuffleNet are presented in Figure 6. The framework employs pointwise group convolutions and a unique "channel shuffle" operation to enhance computational speed while maintaining network efficiency. By utilizing these shuffled pathways, the architecture significantly reduces model size and computational overhead without compromising classification accuracy, making it highly effective for real-time applications.

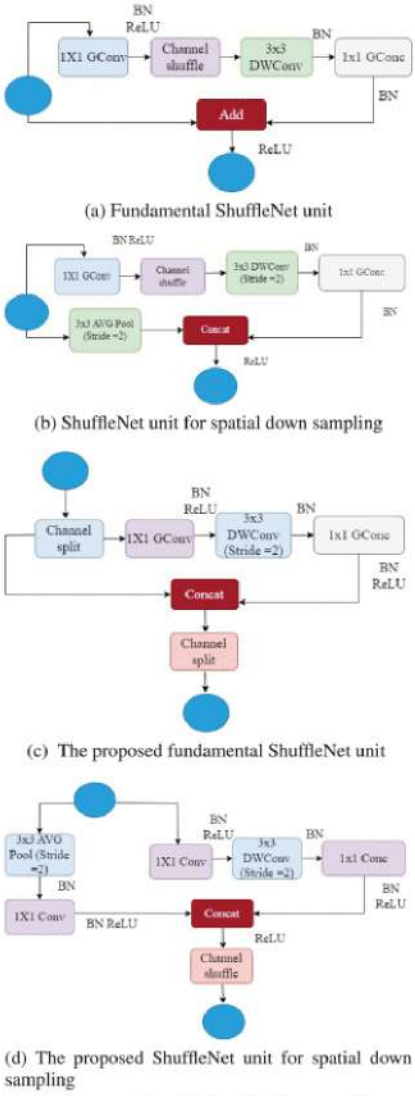


Figure 6. Pre-trained ShuffleNet architecture.

2.2 Genetic Algorithm-Based Feature Selection

The Genetic Algorithm (GA) serves as a metaheuristic optimization technique frequently utilized for feature selection within deep learning frameworks. This methodology is designed to emulate natural evolutionary processes—specifically selection, mutation, and crossover—to identify the most robust and discriminative features while simultaneously minimizing the dimensionality of the feature space. The initial phase of the proposed system involves the generation of a population through random uniform distribution. Within this context, chromosomes are defined as gene-based features, where each constituent gene is assigned a specific real-valued coefficient. The structural representation of these chromosomes is formulated according to the principles established in prior research [27, 28, 29].

2.3 Conventional LSTM

Similar to a Recurrent Neural Network (RNN), a standard Long Short-Term Memory (LSTM) network is engineered to analyze and retain sequential information as it propagates through the architecture. The core of an LSTM consists of various gates and cell states; these components enable the model to selectively ignore or preserve data based on the intensity of cell activity [30]. Specialized neural layers function as gates, regulating both the type of information integrated into the cell state and the specific manner of its application during processing. The cell state acts as a dedicated memory channel, facilitating the transmission of pertinent data across the sequence. Through the training process, the gates dynamically learn to distinguish between critical information and redundant noise.

Communication within the LSTM cell is mediated by input, output, and forget gates. Specifically, these gates determine which data from the previous state $(t-1)$ should be retained, which new information should be added to the current state (t), and which elements should be passed forward to the subsequent step. The architectural configuration of a typical LSTM is illustrated in Figure 7, while the mathematical relationship between inputs and outputs at time t and $t-1$ is defined by the following equations:

$$\zeta_\iota = \vartheta\left[\left(\varsigma_{\rho\zeta} * \rho_\iota\right) + \left(\varsigma_{\upsilon\zeta} * \upsilon_{\iota-1}\right) + \left(\varsigma_{\nu\zeta} * \nu_{\iota-1}\right) + \kappa_\zeta\right] \quad (1)$$

$$\eta_\iota = \vartheta\left[\left(\varsigma_{\rho\eta} * \rho_\iota\right) + \left(\varsigma_{\upsilon\eta} * \upsilon_{\iota-1}\right) + \left(\varsigma_{\nu\eta} * \nu_{\iota-1}\right) + \kappa_\eta\right] \quad (2)$$

$$\gamma_\iota = \left(\eta_\iota * \gamma_{\iota-1}\right) + \zeta_\iota \tanh\left[\left(\varsigma_{\upsilon\gamma} * \upsilon_{\iota-1}\right) + \left(\varsigma_{\nu\gamma} * \nu_{\iota-1}\right) + \kappa_\gamma\right] \quad (3)$$

$$\phi_\iota = \vartheta\left[\left(\varsigma_{\rho\phi} * \rho_\iota\right) + \left(\varsigma_{\upsilon\phi} * \upsilon_{\iota-1}\right) + \left(\varsigma_{\nu\phi} * \nu_{\iota-1}\right) + \kappa_\phi\right] \quad (4)$$

$$\sigma_\iota = \phi_\iota \tanh(\sigma_\iota) \quad (5)$$

where the input gate is represented by $\zeta\iota$, the input vector is denoted by ϱ, the output gate is φ, the output is $\nu\iota$, and the forgetting function is represented by $\eta\iota$. Cell state is expressed as $\gamma\iota$, where γ and κ define bias and weight, respectively.

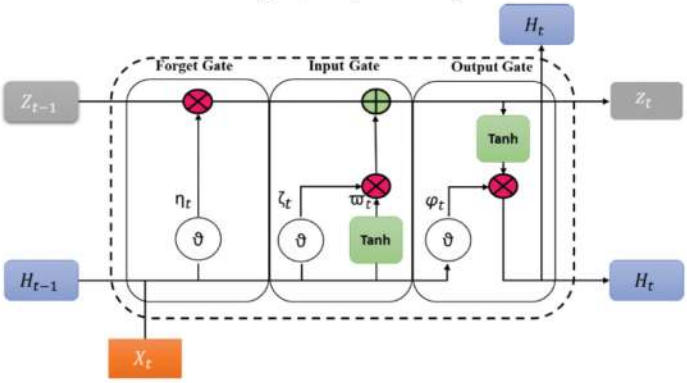


Figure 7. LSTM memory cell.

2.3.1 Bidirectional LSTM (BiLSTM)

The utilization of Bidirectional LSTMs (BiLSTMs) as an enhancement to conventional LSTM architectures significantly improves classification performance by processing data in two temporal directions. In this configuration, two independent LSTMs are trained on the input sequence: one processes the original data stream, while the second processes the sequence in reverse order. This dual-directional approach enriches the network's contextual understanding, enabling it to obtain results more efficiently. Conceptually, the BiLSTM operation is straightforward; it involves duplicating the recurrent layer, where the input is fed forward in its standard sequence through the primary layer and in reverse order through the mirrored layer. This architecture effectively addresses the vanishing gradient problems [31] commonly encountered in traditional RNNs.

The BiLSTM is trained over a temporal span using the entirety of available past and present data. By integrating both forward and backward shifts (scanning from left to right and right to left), the model captures more comprehensive dependencies. In practical implementation, such as with Python's Keras library, the BiLSTM is deployed as a bidirectional wrapper that encapsulates a standard LSTM layer. The resulting combined output from both directions is then propagated to subsequent layers based on a specified fusion mode, as illustrated in Figure 8. The final hidden state is derived from the forward hidden layer $\overrightarrow{\lambda_\tau}$ and the backward hidden layer $\overleftarrow{\lambda_\tau}$ according to the following relationship [32]:

$$\overrightarrow{\Sigma_\iota} = L\left(\varsigma_{\varrho\vec{l}}\varrho_\iota + \varsigma_{\vec{l}\vec{l}}\overrightarrow{l_{\iota-1}} + \kappa_{\vec{l}}\right) \tag{6}$$

$$\overleftarrow{\Sigma_\iota} = L\left(\varsigma_{\varrho\overleftarrow{l}}\varrho_\iota + \varsigma_{\overleftarrow{l}\overleftarrow{l}}\overleftarrow{l_{\iota-1}} + \kappa_{\overleftarrow{l}}\right) \tag{7}$$

$$Z_\iota = \varsigma_{\vec{l}z}\overrightarrow{l_\iota} + \varsigma_{\overleftarrow{l}z}\overleftarrow{l_\iota} + \kappa_z \tag{8}$$

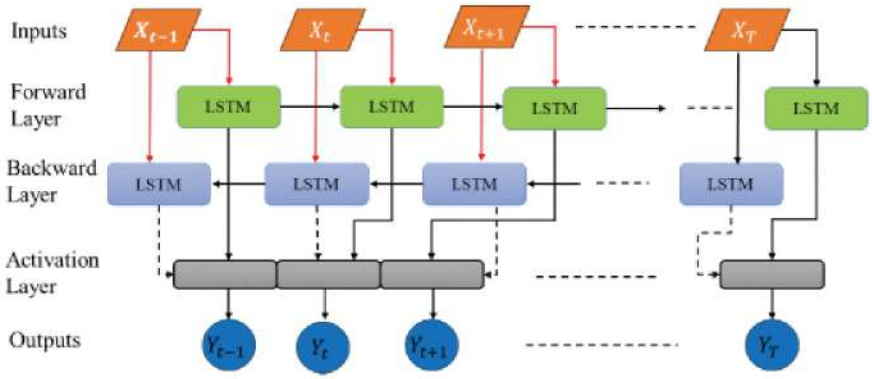


Figure 8. BiLSTM architecture.

In this mathematical framework, the hidden layer is denoted by l, while the matrices representing the input weights for both forward and backward directions are represented by ς. Furthermore, the bias vectors associated with the hidden weights for both directional passes are defined by κ, specifically categorized as $\kappa_{\vec{l}}$ for the forward direction and $\kappa_{\overleftarrow{l}}$ for the backward direction.

4. | RESULTS AND DISCUSSION

This section elucidates the datasets utilized in the current research and the experimental evaluations conducted to demonstrate the efficacy of the proposed model. Statistical validation

of machine learning architectures often involves techniques such as holdout, k-fold cross-validation, and leave-one-out (LOO) cross-validation. However, a significant limitation of the holdout method on constrained datasets is that the training subset may not accurately represent the broader data distribution, potentially leading to biased results. Consequently, cross-validation is regarded as the most robust validation framework when data availability is limited [33, 34, 35]. This approach mitigates the risk of overfitting and facilitates a more precise estimation of the model's predictive performance [36, 37, 38]. Aligning with established research methodologies, we implemented a ten-fold cross-validation scheme and utilized accuracy, precision, recall, and F1-score as primary metrics for the classification analysis.

4.1 Dataset Details

The primary dataset for this investigation was sourced from Kaggle [39]. The research objective was to analyze critical stroke risk factors, including age, gender, smoking status, diabetes, and hypertension. The cohort comprised 100 stroke patients aged 16 and above, with a gender distribution of 68% male and 32% female. The dataset included brain CT imagery categorized as either normal or stroke-affected, encompassing 1,551 normal and 950 stroke images. Notably, the normal images were acquired from the same patients prior to the onset of a stroke. These original grayscale images had a resolution of 650×650 pixels. To prevent overfitting during the training phase, the dataset was randomly balanced, resulting in a finalized set of 950 normal and 950 stroke images, each resized to a resolution of 227×227 pixels.

4.2 Evaluation Metric

In this study, accuracy, precision, recall, and the F1-score served as the fundamental evaluation parameters [40, 41, 42, 43]. These metrics are systematically integrated into all performance reports to ensure a comprehensive assessment. The formal definitions for the utilized evaluation metrics are provided as follows:

$$Accuracy = \frac{No.\,Correct\,Prediction}{Total\,Samples} \tag{9}$$

$$Precision = \frac{TP}{TP+FP} \tag{10}$$

$$Recall = \frac{TP}{TP+FN} \tag{11}$$

$$F1_score = \frac{2TP}{2TP+FP+FN} \tag{12}$$

In this framework, TP represents the count of true positives, while FP denotes the number of false positives. Conversely, TN refers to the total of true negatives, and FN signifies the number of false negatives recorded during the evaluation.

The experimental environment utilized for this study was the Windows 10 operating system, supported by hardware specifications including 32 GB of RAM and an Intel Core™ i7-8250U CPU operating at a frequency of 2 GHz. Software implementation was executed using the Python version 3.7.0 package, deployed through the Anaconda Spyder integrated development environment (IDE).

4.3 Experimental Results and Discussion

This section provides a comprehensive summary of the experimental findings and identifies the baseline architectures utilized for comparative analysis with the proposed model. To rigorously evaluate the effectiveness of our framework, we benchmarked its performance

against a diverse array of Machine Learning (ML) classifiers and Deep Learning (DL) methodologies. For all algorithms involved in this comparative study, default hyperparameters were maintained to ensure a standardized evaluation environment. The empirical results and performance metrics for all implemented ML and DL classifiers are detailed in Table 1 and Table 2, respectively.

Table 1. ML Models Performance

ML models	Accuracy	Precision	Recall	F1-score
Logistic Regression	75.62	68.56	78.40	75.00
Decision Tree	85.00	80.65	88.34	85.00
Random Forest	89.53	91.50	86.73	90.00
Naive Bayes	66.21	70.00	62.15	61.00
Support Vector Machine	91.45	89.88	92.00	91.00

Table 2. DL Models Performance

DL models	Accuracy	Precision	Recall	F1-score
ANN	92.50	88.56	91.40	92.00
DNN	95.78	95.26	88.62	95.00
RNN	93.66	91.50	89.00	94.00

Based on the comparative analysis presented in Table 1, the Support Vector Machine (SVM) demonstrates superior diagnostic performance among all traditional machine learning classifiers, achieving a peak accuracy of 91.45% on the evaluated dataset. Conversely, the Naive Bayes classifier exhibits the lowest efficacy, yielding a diagnostic accuracy of only 66.21%. Regarding the deep learning architectures summarized in Table 2, the Deep Neural Network (DNN) attains the highest accuracy at 95.78%, notably surpassing the performance of both the Artificial Neural Network (ANN) at 92.50% and the Recurrent Neural Network (RNN) at 93.66%.

To address challenges associated with imbalanced data distribution and restricted sample sizes, we conducted evaluations of the proposed model, wherein features identified by the neural network-based genetic algorithm were integrated into LSTM and BiLSTM frameworks for stroke identification. Two subsequent experiments were executed utilizing the GA-based feature extraction process on brain imagery. In the primary experiment, an LSTM was deployed for the classification of stroke pathology. Feature selection was performed by establishing a specific threshold following a comprehensive ranking of traits, after which the most discriminative features were channeled into the LSTM architecture. To rigorously assess the GA_LSTM model's diagnostic capability, a 5-fold cross-validation scheme was implemented, utilizing the Receiver Operating Characteristic (ROC) curve as the primary evaluative tool [35, 44, 45, 46]. As illustrated in Figure 9, the GA_LSTM configuration achieved a significant Area Under the Curve (AUC) of 95%.

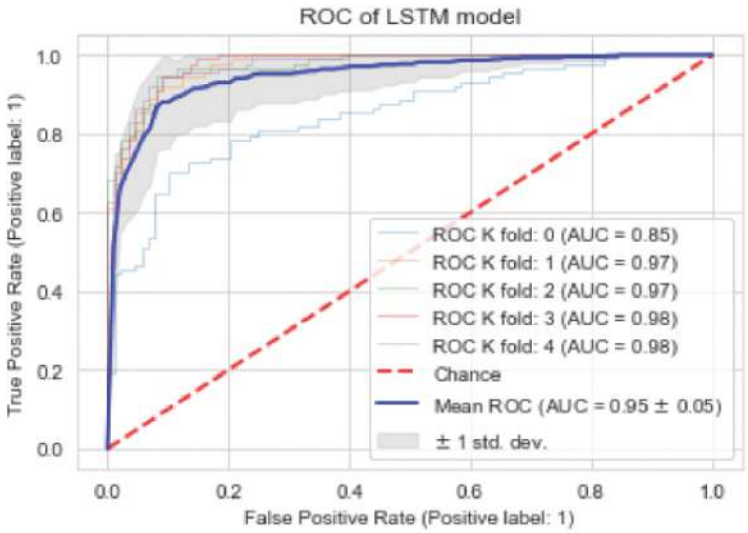


Figure 9. ROC of GA_LSTM.

In the subsequent phase of experimentation, the BiLSTM model was integrated with the primary features identified by the neural network-based genetic algorithm; consequently, this proposed architecture was designated as the GA_BiLSTM model. As demonstrated in Figure 10, it is clear that the GA_BiLSTM framework achieved the superior Area Under the Curve (AUC) on the ROC plot, outperforming the previous configurations. Furthermore, an exhaustive evaluation of the accuracy performance of the newly developed GA_BiLSTM model was conducted, the results of which are explicitly depicted in Figure 11.

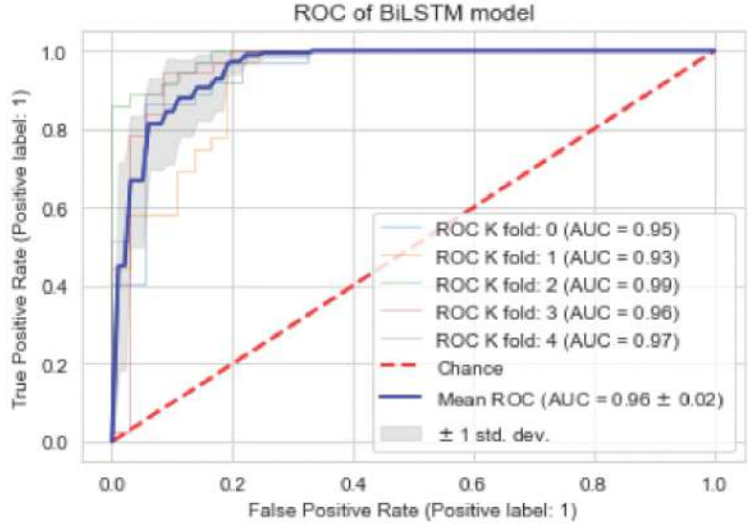


Figure 10. ROC of GA_BiLSTM.

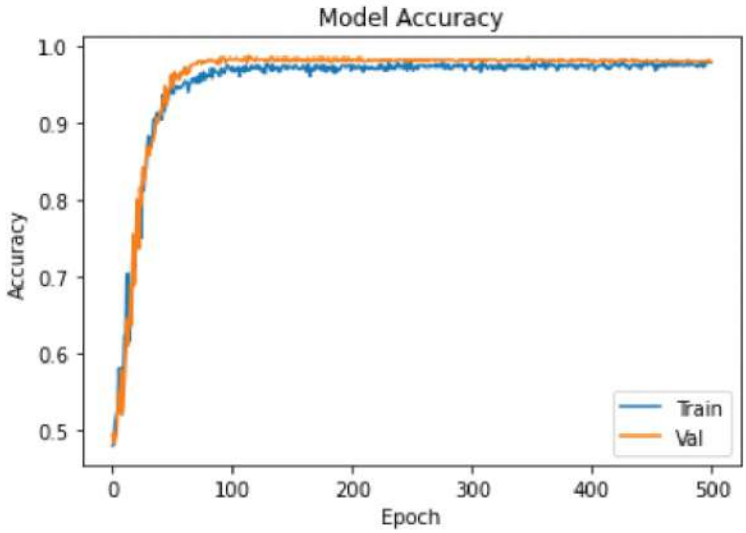


Figure 11. Accuracy curves of GA_BiLSTM.

Furthermore, the performance of the GA_BiLSTM architecture was rigorously evaluated alongside the GA_LSTM model based on the metrics of accuracy, precision, recall, and F1-score, as detailed in Table 3. Analysis of the data in Table 3 reveals that the proposed GA_BiLSTM model exhibits superior performance compared to both the GA_LSTM framework and the various deep learning and machine learning models previously documented in Table 1 and Table 2. This indicates that the integration of bidirectional temporal processing with genetic algorithm-based feature extraction provides a more robust solution for accurate stroke detection.

Table 3. Proposed Models Performance

Models	Accuracy	Precision	Recall	F1-score
GA_LSTM	95.35	92.00	90.89	95.00
GA_BiLSTM	96.45	98.00	93.50	96.00

4.4 Previously Proposed Methods for Stroke Prediction

In this section, the performance of the newly introduced stroke detection framework is evaluated and benchmarked against various machine learning methodologies previously proposed by the research community. To ensure a rigorous comparison, the classification accuracy of the proposed model was analyzed alongside other studies that utilized the identical Kaggle dataset of brain CT imagery for stroke prediction. The comparative analysis of accuracy between the proposed GA_BiLSTM architecture and existing methodologies is summarized in Table 4. It is important to highlight that this newly developed method demonstrates a substantial enhancement in diagnostic accuracy over prior investigations that employed conventional machine learning and deep learning approaches for stroke identification.

Table 4. Comparison with Existing Studies

Study (Year)	Method	Accuracy(%)
Sailasya et.al., (2021) [47]	LR	78.00
	DT	66.00
	RF	73.00
Aniwat et.al., (2021) [17]	DNN	96.21
Devet et.al., (2022) [15]	SVM	68.00
	RF	74.00
	CNN	74.00
Santwana et.al., (2022) [18]	RF	87.97
Akter et.al., (2022) [48]	RF	95.30
Tursynova et.al., (2023) [49]	CNN	81.00
Santwana et.al., (2023) [19]	RF	87.22
Yeo et.al., (2023) [20]	RNN + CNN	93.00
Gupta et.al., (2023) [50]	DenseNet-121	96.00
Luis et.al., (2023) [51]	DNN	92.70
Current Study	GA + LSTM	95.35
	GA + BiLSTM	96.45

5. | CONCLUSION

This study introduces a robust methodology for stroke detection by leveraging advanced machine learning and deep learning techniques. The core of the proposed framework is the GA_BiLSTM model, which utilizes a neural network-based genetic algorithm to identify and extract critical features from brain CT imagery. These refined features are subsequently processed through LSTM and BiLSTM architectures to facilitate accurate stroke prediction. To validate the model's efficacy, various K-fold cross-validation schemes and alternative machine learning algorithms were benchmarked. The experimental findings demonstrate that the proposed model significantly outperforms existing baseline classifiers, achieving a peak accuracy of 96.45%.

The integration of Genetic Algorithms for feature selection and BiLSTM for temporal processing proved more efficient than traditional ML and standard DL models. However, while the model performed exceptionally well on the current dataset, it is acknowledged that deep learning models typically yield even better results when trained on larger, more diverse datasets. Furthermore, performance remains highly dependent on the quality of the input data, necessitating future improvements in data preprocessing and acquisition.

Looking forward, research will focus on collecting larger sample sizes to further enhance the model's generalization and predictive accuracy. We aim to implement more sophisticated architectures capable of fully autonomous stroke prediction and develop novel methods to improve the quality and consistency of medical imaging data. Additionally, to foster trust within the medical community, future work will prioritize the integration of explainable AI (XAI) techniques. These methods will clarify the model's internal decision-making process, ultimately assisting healthcare professionals in making more informed and transparent clinical decisions based on AI-generated results.

REFERENCES

1. World Stroke Organization. (2023). Vision Impairment and Blindness. Accessed: Feb. 3, 2023. [Online]. Available: world-stroke.org.
2. A. Khosla, et al., "An integrated machine learning approach to stroke prediction," in *Proc. 16th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Jul. 2010, pp. 183–192.
3. Centers for Disease Control Prevention. (2023). Stroke Facts. Accessed: Mar. 25, 2023. [Online]. Available: cdc.gov.
4. Y. Xu, et al., "Role of ferroptosis in stroke," *Cellular Mol. Neurobiol.*, vol. 43, no. 1, pp. 205–222, Jan. 2023.
5. X. Li and G. Chen, "CNS-peripheral immune interactions in hemorrhagic stroke," *J. Cerebral Blood Flow Metabolism*, vol. 43, no. 2, pp. 185–197, Feb. 2023.
6. Y. Yang, et al., "Accurate prediction of stroke for hypertensive patients based on medical big data and machine learning algorithms: Retrospective study," *JMIR Med. Informat.*, vol. 9, no. 11, Nov. 2021, Art. no. e30277.
7. M. Chun, et al., "Stroke risk prediction using machine learning: A prospective cohort study of 0.5 million Chinese adults," *J. Amer. Med. Inform. Assoc.*, vol. 28, no. 8, pp. 1719–1727, Jul. 2021.
8. G. Ntaios, et al., "Machine-learning-derived model for the stratification of cardiovascular risk in patients with ischemic stroke," *J. Stroke Cerebrovascular Diseases*, vol. 30, no. 10, Oct. 2021, Art. no. 106018.
9. K. Uchida, et al., "Development of machine learning models to predict probabilities and types of stroke at prehospital stage: The Japan urgent stroke triage score using machine learning (JUST-ML)," *Transl. Stroke Res.*, vol. 13, no. 3, pp. 370–381, Jun. 2022.
10. H. Z. U. Rehman, et al., "Conventional and deep learning methods for skull stripping in brain MRI," *Appl. Sci.*, vol. 10, no. 5, p. 1773, Mar. 2020.
11. M. Lim, et al., "Link prediction in time-evolving criminal network with deep reinforcement learning technique," *IEEE Access*, vol. 7, pp. 184797–184807, 2019.
12. M. Lim, et al., "Hidden link prediction in criminal networks using the deep reinforcement learning technique," *Computers*, vol. 8, no. 1, p. 8, Jan. 2019.
13. S. Bacchi, et al., "Deep learning in the prediction of ischaemic stroke thrombolysis functional outcomes," *Academic Radiol.*, vol. 27, no. 2, pp. e19–e23, Feb. 2020.
14. M. Saritha, et al., "Classification of MRI brain images using combined wavelet entropy based spider web plots and probabilistic neural network," *Pattern Recognit. Lett.*, vol. 34, no. 16, pp. 2151–2156, Dec. 2013.
15. S. Dev, et al., "A predictive analytics approach for stroke prediction using machine learning and neural networks," *Healthcare Anal.*, vol. 2, Nov. 2022, Art. no. 100032.
16. L. Ali, et al., "An automated diagnostic system for heart disease prediction based on chi square statistical model and optimally configured deep neural network," *IEEE Access*, vol. 7, pp. 34938–34945, 2019.
17. A. Phaphuangwittayakul, et al., "An optimal deep learning framework for multi-type hemorrhagic lesions detection and quantification in head CT images for traumatic brain injury," *Int. J. Speech Technol.*, vol. 52, no. 7, pp. 7320–7338, May 2022.
18. S. Gudadhe and A. Thakare, "Classification of intracranial hemorrhage CT images for stroke analysis with transformed and image-based GLCM features," *Tech. Rep.*, 2022.

19. S. Gudadhe, et al., "A novel machine learning-based feature extraction method for classifying intracranial hemorrhage computed tomography images," *Healthcare Anal.*, vol. 3, Nov. 2023, Art. no. 100196.
20. M. Yeo, et al., "Evaluation of techniques to improve a deep learning algorithm for the automatic detection of intracranial haemorrhage on CT head imaging," *Eur. Radiol. Experim.*, vol. 7, no. 1, p. 17, Apr. 2023.
21. K. Khosravi, et al., "Soil water erosion susceptibility assessment using deep learning algorithms," *J. Hydrol.*, vol. 618, Mar. 2023, Art. no. 129229.
22. J. Donnelly, et al., "Gaussian process emulation of spatio-temporal outputs of a 2D inland flood model," *Water Res.*, vol. 225, Oct. 2022, Art. no. 119100.
23. E. Rasool, et al., "Breast microcalcification detection in digital mammograms using deep transfer learning approaches," in *Proc. 9th Int. Conf. Comput. Data Eng.*, Jan. 2023, pp. 58–65.
24. J. Donnelly, et al., "Forecasting global climate drivers using Gaussian processes and convolutional autoencoders," *Eng. Appl. Artif. Intell.*, vol. 128, Feb. 2024, Art. no. 107536.
25. M. A. Saleem, et al., "Sooty tern optimization algorithm-based deep learning model for diagnosing NSCLC tumours," *Sensors*, vol. 23, no. 4, p. 2147, Feb. 2023.
26. B. Ghiasi, et al., "Uncertainty quantification of granular computing-neural network model for prediction of pollutant longitudinal dispersion coefficient in aquatic streams," *Sci. Rep.*, vol. 12, no. 1, p. 4610, Mar. 2022.
27. S. Sindhiya and S. Gunasundari, "A survey on genetic algorithm based feature selection for disease diagnosis system," in *Proc. IEEE Int. Conf. Comput. Commun. Syst.*, Feb. 2014, pp. 164–169.
28. A. Javeed, et al., "Predicting dementia risk factors based on feature selection and neural networks," *Comput., Mater. Continua*, vol. 75, no. 2, pp. 2491–2508, 2023.
29. P. Alirezazadeh, et al., "A genetic algorithm-based feature selection for kinship verification," *IEEE Signal Process. Lett.*, vol. 22, no. 12, pp. 2459–2463, Dec. 2015.
30. R. Noori, et al., "An efficient data driven-based model for prediction of the total sediment load in rivers," *Hydrology*, vol. 9, no. 2, p. 36, Feb. 2022.
31. A. Javeed, et al., "Predictive power of XGBoost_BiLSTM model: A machine-learning approach for accurate sleep apnea detection using electronic health data," *Int. J. Comput. Intell. Syst.*, vol. 16, no. 1, p. 188, Nov. 2023.
32. A. Mousa and B. Schuller, "Contextual bidirectional long short-term memory recurrent neural network language models: A generative approach to sentiment analysis," *Tech. Rep.*, 2017.
33. A. Javeed, et al., "Early prediction of dementia using feature extraction battery (FEB) and optimized support vector machine (SVM) for classification," *Biomedicines*, vol. 11, no. 2, p. 439, Feb. 2023.
34. S. Yadav and S. Shukla, "Analysis of k-Fold cross-validation over hold-out validation on colossal datasets for quality classification," in *Proc. IEEE 6th Int. Conf. Adv. Comput. (IACC)*, Feb. 2016, pp. 78–83.
35. A. Javeed, et al., "Breaking barriers: A statistical and machine learning-based hybrid system for predicting dementia," *Frontiers Bioeng. Biotechnol.*, vol. 11, Jan. 2024, Art. no. 1336255.
36. A. Javeed, et al., "An intelligent learning system for unbiased prediction of dementia based on autoencoder and AdaBoost ensemble learning," *Life*, vol. 12, no. 7, p. 1097, Jul. 2022.

37. K. Pal and B. V. Patel, "Data classification with k-fold cross validation and holdout accuracy estimation methods with 5 different machine learning techniques," in *Proc. 4th Int. Conf. Comput. Methodologies Commun. (ICCMC)*, Mar. 2020, pp. 83–87.
38. A. Javeed, et al., "A clinical decision support system (CDSS) for unbiased prediction of caesarean section based on features extraction and optimized classification," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–13, Jun. 2022.
39. (2021). Brain Stroke CT Image Dataset. Accessed: Jul. 3, 2023. [Online]. Available: kaggle.com.
40. J. O. Westgard and J. A. Lott, "Precision and accuracy: Concepts and assessment by method evaluation testing," *CRC Crit. Rev. Clin. Lab. Sci.*, vol. 13, no. 4, pp. 283–330, Jan. 1981.
41. H. Dalianis, "Evaluation metrics and evaluation," in *Clinical Text Mining: Secondary Use of Electronic Patient Records*, 2018, pp. 45–53.
42. L. Ali, et al., "Parkinson's disease detection based on features refinement through L1 regularized SVM and deep neural network," *Sci. Rep.*, vol. 14, no. 1, p. 1333, Jan. 2024.
43. A. Javeed, et al., "Decision support system for predicting mortality in cardiac patients based on machine learning," *Appl. Sci.*, vol. 13, no. 8, p. 5188, Apr. 2023.
44. X. Zhang, et al., "The use of ROC and AUC in the validation of objective image fusion evaluation metrics," *Signal Process.*, vol. 115, pp. 38–48, Oct. 2015.
45. L. Ali, et al., "Reliable Parkinson's disease detection by analyzing handwritten drawings: Construction of an unbiased cascaded learning system based on feature selection and adaptive boosting model," *IEEE Access*, vol. 7, pp. 116480–116489, 2019.
46. T. Fawcett, "An introduction to ROC analysis," *Pattern Recognit. Lett.*, vol. 27, no. 8, pp. 861–874, Jun. 2006.
47. 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, 2021.
48. B. Akter, et al., "A machine learning approach to detect the brain stroke disease," in *Proc. 4th Int. Conf. Smart Syst. Inventive Technol. (ICSSIT)*, Jan. 2022, pp. 897–901.
49. A. Tursynova, et al., "Deep learning-enabled brain stroke classification on computed tomography images," *Comput., Mater. Continua*, vol. 75, no. 1, pp. 1431–1446, 2023.
50. M. Gupta, et al., "Predicting brain stroke using IoT-enabled deep learning and machine learning: Advancing sustainable healthcare," in *Proc. Int. Conf. Sustain. Develop. Mach. Learn., AI IoT. Cham, Switzerland: Springer*, 2023, pp. 113–122.
51. L. Cortés-Ferre, et al., "Deep learning applied to intracranial hemorrhage detection," *J. Imag.*, vol. 9, no. 2, p. 37, Feb. 2023.