MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, Jan-June 2026, Pages: 1810-1817

MACHINE LEARNING-BASED EARLY STROKE PREDICTION USING CONVOLUTIONAL NEURAL NETWORKS

[1*]Theophilus F
[1]Assistant Professor, Department of Artificial Intelligence and Data Science,*
V.S.B College of Engineering Technical Campus, Coimbatore -642109, Tamilnadu, India.
[2]N.Vijayakumar
[2]Assistant Professor, Department of Computer science and Technology,
Karpagam College of Engineering,Coimbatore 641105, Tamilnadu, India.
[3]K Madhu Suganya
[3]Assistant Professor, Department of Information Technology,
Nandha college of Technology, Erode - 638052,Tamilnadu, India.
[4]M Vasanthakumar
[4]Assistant Professor, Department of ECE, AVS Engineering College,
Ammapet,Salem-636003, Tamilnadu, India.
[1]Corresponding Author Mail Id: ftheophilus435@gmail.com*

ABSTRACT

A brain stroke is mostly brought on by modifications in the blood flow to certain brain regions. The patient's quality of life may consequently be diminished as a result of some specific functions associated with the affected area being restricted.As a result, machine learningbased prediction approaches used in clinical decision, including anticipating the onset and course of diseases as well as assisting physicians for medication.This approach to stroke predictive analytics technology was implemented using machine learning models on brain disease datasets. The proposed of this model is to create a machine learning application to identify stroke.Low prediction accuracy, precision, and recall, as well as high temporal complexity and detection error rate, are some of the disadvantages of ML approaches. To address these problems, we suggest a novel approach called the machine learning-based Convolutional Neural Network (CNN) model, which can accurately and efficiently predict a patient's risk of stroke. Brain illness datasets were used to test machine learning models for this predictive stroke analysis technique.This model aims to develop a machine learning application that uses CNN to identify strokes. This model, which is an ad hoc version of a multilayer perceptron, is employed in feature selection to determine an attribute's maximum threshold. Because it is a healthcare dataset, the predictive model only uses 11 features and one target class. Consequently, in order to extract the features that have the using feature selection methods. With a 92.5% accuracy rate, the model surpassed other machine learning models in our comparison of their accuracy.

Keywords: *Stroke, Machine Learning, Gated Convolution Neural Network, Classification, Multi-Layer Perceptron, Brain.*

1. INTRODUCTION

Machine learning techniques are increasingly employed to find trends in datasets because annotated medical record databases have become available. Any medical ailment can be accurately diagnosed with the use of this analysis. Enhancing health and cutting down on medical expenses. The biomedical and health care industries have greatly benefited from the use of data mining techniques in medical records. It aids medical professionals in early disease detection. Finding the primary risk factors and contributing variables for stroke is of special importance to experts.

Stroke is an event where irregularities in the brain's blood arteries lead to dysfunction in particular brain regions, in the past 40 years, the occurrence of stroke has more than doubled in emerging economies. Early identification is critical because there is currently no proven treatment for stroke. The most widely used techniques for identifying stroke illness are CT and MRI scans. However, due to their high cost, CT and MRI may not be appropriate for emerging or low-income countries.

Low-cost, accurate, and quick stroke detection technologies are desperately needed by healthcare services because stroke is a major disease globally, particularly among older and low-income populations.

The prediction and early detection of strokes have been extensively studied. Patient and cardiovascular illness rate [4]. According to the findings, cardiovascular disease accounted for 33% of all deaths in 2019. According to the Cardiovascular File, ischemic stroke accounts for 11.1 million of the 18.6 million deaths that occurred in 2019 and causes the majority of cardiovascular events. Furthermore, the chance of an ischemic stroke recurring is extremely high; at two years, the rate has been estimated to be 14.1% [5]. Furthermore, a link between the coronavirus and stroke has been shown in more recent research, increasing the risk of stroke-related mortality.

Blood tests, neurological physiological techniques, such as started expected testing, and cerebrum imaging techniques, including as CT, X-ray, and X-beam imaging, can all be used to diagnose stroke infection. Despite its drawbacks, such as radiation exposure and perhaps adverse reactions from the differentiating specialists, CT and X-ray are the most often used methods for stroke diagnosis. These devices can also be difficult to use because they cost a lot for each evaluation, need small workspaces, and demand careful supervision – all of which make finding them quite difficult. With the help of new wearable cathodes, members can assess cerebrum stroke while relaxing in their own homes. These cathodes, which are attached to the skull, track the motion of the nerve cells in the brain.Additionally, brain impulses are captured during the various sleep stages, enabling quick and painless investigation. Both the patient's movements and background noise can taint BRAIN STROKE data. Nevertheless, compared to the previously described imaging methods, it is feasible to gather and evaluate BRAIN STROKE data in real time, affordably, and with fewer adverse effects. These benefits make 24-hour BRAIN STROKE measures a practical and affordable way to track stroke disease, which has a high recurrence rate in everyday life.

Analysed disorders including stroke, Alzheimer's disease, and seizures using stroke tests; other studies have utilised stroke tests to correlate emotions and sleep states. Nevertheless, the majority of these investigations retrieved predetermined frequency components for straightforward tests or classification. Health status monitoring is not currently possible with the presented methodologies since segmenting the raw data into the frequency domain takes more time and money.A recent study that used solely raw stroke data was able to categorise stroke patients, which is a promising first step towards using stroke for real-time health monitoring.

Created a browsing routine that is indicative of Korean seniors' daily routines. The measured stroke data utilized in this investigation were submitted by Korean people 65 years of age and older. The acquired stroke data were divided into two categories: raw data and frequency domain extracted data, in order to facilitate a comparison between the machine learning and machine learning models. After the Fast Fourier Transform (FFT) was applied to each raw brain stroke data from the six channels (Fz, T7, C1, C2, T8, and Oz), 66 values in total were gathered and used in the experiment. According to preliminary studies, stroke can be accurately predicted using only raw data.The 94.0% precision of the CNN model in our examinations, alongside its 5.7% bogus negative rate (FNR) and 6.0% misleading positive rate (FPR), all help the serious level of trust in the outcomes. By utilizing the power values, nonetheless, the CNN model showed a precision of 81.4%, a FPR of 18.5%, and a FNR of 17.3%. Moreover, we found that in tests using relative qualities, the CNN model had a precision of 89.2%, with FPR and FNR of 12.5% and 8.4%, separately.

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, Jan-June 2026, Pages: 1810-1817

These exploratory outcomes show that the simple stroke information alone can be utilized to dependably recognize and gauge stroke without the need to isolate out the recurrence trait values. Besides, by precisely foreseeing the prognostic side effects of stroke continuously — a condition with very high mortality and repeat rates — the framework portrayed in this work offers a cheap method for checking the wellbeing status of more established individuals in their everyday exercises.The key contributions of this work are summarized below.

- Developing a sophisticated CNN model that, using a benchmark dataset, can forecast stroke.
- Finding the model's optimal features by Using various feature selection algorithms
- Using a rigorous simulation to assess the model's performance against several machine learning models and displaying the comparative outcome based on various performance metrics.

2. LITERATURE SURVEY

Stroke is a condition where irregularities in the brain's blood arteries lead to dysfunction in particular brain regions. An increasing amount of study aims to predict and analyse specific diseases beforehand as a result of advancements in healthcare technology. There is a lot of research being done on stroke disease, especially as the population grows elderly.

Due to its high mortality and recurrence rates, stroke is a disease that needs to be continuously observed and monitored by medical professionals. Stroke is particularly deadly for the elderly [1]. Health-related behaviors are one of the primary risk factors for stroke and are becoming more and more significant as a preventive intervention. Many machine learning models have been created to use variables like lifestyle characteristics and radiographic images to detect stroke or predict the risk of stroke. However, no model has been created with data from laboratory experiments [2].

To enable prompt clinical action to lessen the severity of stroke, early detection of the several warning symptoms of the condition is essential. Robust feature representation skills enable deep neural networks to automatically extract discriminative features from massive data sets [3]. The lives of stroke survivors are drastically affected, and disability is often the outcome. Manual stroke analysis is laborious and subject to operator variances [4]. Thanks to the increasing convergence of clinical diagnoses and technology, caregivers can now effectively manage their patients by meticulously locating and maintaining patient medical records. Thus, it's crucial to evaluate how these risk factors interact with one another in the patient's medical records and ascertain how much each risk factor contributes in relation to the chance of having a stroke [5].

Both patients and physicians value the ability to forecast functional outcome after an ischemic stroke (IS). This makes it easier for medical professionals to collaborate with patients and their families in an efficient manner, enabling them to set realistic objectives and decide together on post-stroke rehabilitation treatments and exercise plans that will promote healing [6]. On the Kaggle website, there is a reliable dataset for stroke prediction that can be used to evaluate the algorithm's performance. Ada Boost Deep Neural Networks (3-layer and 4-layer ANN), Random Forest, Decision Tree, Logistic Regression, K-Nearest Neighbors, SVM - Linear Kernel, Naive Bayesian, Ada Boost Extreme Gradient Boosting (XGBoost), and Light Gradient Boosting Machine [7].

Missing and imbalanced data make it very difficult to find hidden risk variables and make reliable forecasts. Three different imputation techniques were used in this investigation to handle missing data [8]. Early detection and recognition of various stroke warning signs May reduce stroke severity. Several machine learning (ML) techniques have been developed to estimate stroke risk.

Stroke prognosis is dependent on numerous clinical and patient-specific factors, accurate outcome prediction is still challenging. ML has a lot to offer in the diagnosis and prediction of strokes, but it's critical to recognise that the effectiveness of the process depends on the calibre of the input data and how well the algorithm works with a given collection of data. Furthermore, targeted data gathering initiatives could improve upcoming assessments of machine learning techniques for stroke. Due to the reality that small sample sizes often constrain machine learning research [9]. Because the prognosis of stroke is dependent on numerous clinical and patient-specific variables, accurate outcome prediction stays difficult [10].

It can be challenging to forecast future short- or long-term care demands and to create realistic treatment goals because different patients have varying possibilities for recovery. This study set out to assess the prognostic value of high-resolution data on poststroke gait function acquired from wearable wireless motion sensors after inpatient stroke treatment [11].

Predicting the prognosis of stroke patients is critical. An interpretable machine learning model for estimating the likelihood of a stroke-related mortality within a year. Our sub module emphasizes the temporal and spatial variability of various variables while attempting to retrieve features from raw clinical data. Developed a novel relational attention module that uses bidirectional long short-term memory (Bi-LSTM) as the model and takes variable correlation into account [12]. It is among the world's main causes of death and long-term problems.

Stroke risk prediction systems that are currently in use can be defeated by machine learning-based methods [13]. The majority of the thousands of components that are frequently present in electronic medical records are superfluous or redundant and should be eliminated in order to increase prediction accuracy [14]. Fault prediction is a crucial assurance for enhancing system dependability in electromagnetic radiation (EML) systems. In spite of this, no fully developed technology is now usable directly [15]. Manual disease classification is laborious, prone to mistakes, and unable to handle data sets [16].In research including healthy subjects, electroencephalography (EEG)-based BCI was utilised to ascertain the direction of hypothetical hand movements. Using EEG recordings, we investigated MI-based brain activity to ascertain the direction of imagined hand movements in stroke patients [17].

Applying machine learning models to disease prediction. A model was developed [18] that connected convolutional neural networks with long short-term memory. However, as interpretations can give medical practitioners insight into the factors influencing risk levels, they might be more significant in this regard than accuracy. Data loss typically occurs in medical information records [19]. Two crucial metrics for evaluating the efficacy of training and programme recommendations are the 6-minute walk distance (6MWD) for stroke patients and the Fugl-Meyer Lower Extremity Assessment Scale (FMA-LE).

However, as mobility assessment in stroke patients is usually done by manual observation and table scoring, issues with lost labor and subjective observational findings occur [20]. The current Thrombolysis in Cerebral Infarction (TICI) score varies between and among observers because it is based on an approximate ordinal grade of ocular assessment. In this study, we introduce AutoTICI, a quantitative automatic TICI scoring method [21].

Manual segmentation is frequently an arbitrary, time-consuming procedure. Depending on how long it takes to remove the blood clot, patients with acute ischemic stroke have a wide range of outcomes [22]. As soon as you've determined and assessed risk, get prompt, efficient therapy. Consequently, there is substantial therapeutic value in researching methods for the automatic identification and classification of stroke lesions [23]. Finding ischemic areas in AIS patients that contain both possibly recoverable tissue (penumbra) and dead tissue (core) may help with diagnosis in the following ways. Quick and precise prediction technique. Because it can offer useful data for therapy planning, it has a big influence on treatment.

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, Jan-June 2026, Pages: 1810-1817

When a stroke occurs, computed tomography perfusion (CTP) is frequently the main technique used to identify the location, volume, and severity of ischemic regions. The automated segmentation methods for CTP that are now on the market usually make use of preprocessed 3D parametric maps that radiologists employ for clinical interpretation. Nevertheless, rather than finding stroke predictors, prior research has primarily concentrated on creating acute treatment plans and clinical treatment guidelines following the onset of stroke [25].

Brain tissue damage caused by cerebral ischemic stroke is typically characterized by a complex series of neuronal and vascular processes. Early developmental imaging can identify optimal treatment options to restore functional brain damage and predict tissue infarction and margins. Therefore, new imaging methods capable of identifying brain damage in the early stages of ischemic stroke are urgently needed [26].

Axial nerve depicts were flipped on the x-axis using the contralateral-ipsilateral hemispheric mirror (MCI) approach in order to estimate prognosis. All agree, nonetheless, that the brain's left and right hemispheres are asymmetrical. MCI mixes together asymmetries caused by unilateral injury with innate brain asymmetries, which creates problems with physiological interpretation and raises questions about the validity of findings [27].

After a stroke, the dynamic rearrangement of functional brain networks is still not well understood. In this investigation, classified 15 stroke patients into two subgroups based on clinical symptoms: mild (n = 6) and severe (n = 9) using resting-state functional MRI data [28]. One of the biggest challenges in stroke care is specialised rehabilitation [29].

An increasing number of people are at risk of stroke as the population ages, which emphasises the need for precise, trustworthy prediction systems. Six well-known classifiers were used in a comparison analysis to evaluate the effectiveness of the suggested machine learning technique based on criteria for prediction accuracy and generalisation capacity [30]. Non-uniform scan volumes provide challenges to computer-based parallel assessment algorithms, necessitating radiologists to manually choose the most significant slices [31].

The spatiospectral features that most effectively identify stroke survivor SN participants using EEG responses to ipsilateral and contralateral visual stimuli were identified using an electroencephalogram (EEG)-based brain-computer interface (BCI). [32]. But only some areas of the cerebrovasculature can be reached due to the limited access provided by the cranial acoustic window [33]. In the context of microwave imaging, it was thought that the field of predictive validation with spatial interpretation could help with the unreliability brought on by these models' "black box" nature [34]. Using a new imaging sequence that combines sparse sampling with fast trajectory, we produced Whole-brain map (2.0 × 3.0 × 3.0 mm3) and quantitative T2 values (1.9 × 1.9 × 3.0 mm3) of neurometabolites [35].

However, limb-specific classification performance remains low and does not account for patient differences resulting from brain reorganization after stroke. EEG data from 11 stroke patients and 11 healthy controls were used [36] . The robotic redundant limb SoftHandThis would allow the hand to be used again, partially restore functioning, and motivate people to avoid acquired disuse.

To predict the possibility of a different treatment's outcome, apply a machine learning predictive model developed for the first treatment [38]. Transcutaneous vagus nerve stimulation (tVNS) improved both ipsilateral and contralateral upper limb motor and cognitive function in stroke patients after just one session [39]. Stroke or brain disease is one of the main causes of disability in adults and humans. Seeking assistance is essential as this is a medical emergency [40]. Following a stroke, people' recovery times vary depending on their symptoms and level of organ damage. Patients who recognise their risk and take proactive steps to lower it can avoid up to 80% of stroke [41].

Electronic health data, however, are less helpful in researching the interactions between various stroke risk factors [42]. Finding relevant information on a medical condition on the internet can be difficult at times. More definitive research on the relationship between social networks and stroke risk is expected to improve the precision with which working persons can identify stroke risk [43]. The clinical intricacy of post-stroke damage makes it difficult to anticipate upper limb (UL) functional recovery during therapy, even though stroke is one of the most common causes of disability [44]. The choices a doctor takes about treatment during the acute stage of a stroke have a big influence on the patient's prognosis. Over the past five years, Many ratings, including ASTRAL, DRAGON, and THRIVE, have been proposed as tools to help physicians predict functional outcomes in stroke patients [45].

A practical method for predicting intrinsic stroke (ICT) in stroke patients. Twelve chronic stroke patients who used a treadmill to demonstrate their ICT had their gait kinematics examined. The outlier principle-based prediction algorithm is used. To identify antecedents for precise ICT prediction, Sequential Forward Selection (SFS) and minimum redundancy maximum correlation (mRMR) were applied, respectively [46].With the growing synergy between medical diagnostics and technology, caregivers can improve patient management by carefully mining and preserving patients' medical records. As a result, it's critical to look at how these risk factors interact in patient medical records and ascertain how each risk factor contributes to the likelihood of stroke [47]. MRI using CNN and machine learning algorithms to diagnose strokes [48]. Semantic segmentation is used to distinguish the abnormal regions in stroke MRI images after the pictures have been classified as normal or abnormal [49]. Medical image analysis is best suited for models built on Convolutional Neural Networks (CNNs). The rate at which stroke is occurring is concerning, and improved techniques are required to identify strokes promptly [50].

3. MATERIALS AND METHOD

As illustrated in Figure 1, we suggest a unique machine learning-based stroke prediction system that makes use of the original stroke values and attribute values acquired in real time. The following modules make up the suggested system. (1) Module for collecting data. (2) Module for preparing data. (3) Module for biological data analysis. (4) A machine learning-based learning and prediction module. (5) Analysis of stroke prediction based on biological cues. In feature selection, biosignal-based modules extract and manage frequency properties. Next, a series of preprocessing actions and features gathered from various biological data, including older adult stroke data, are gathered and stored by the proprietary system. In order to predict and evaluate stroke, a machine learning model is lastly constructed using the biosignal data that has been collected.

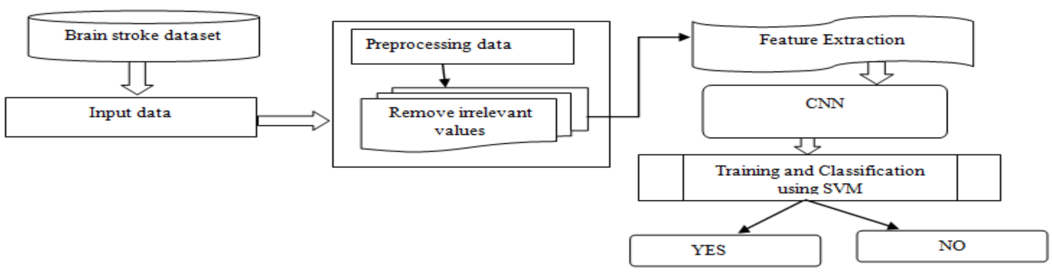

Figure 1: Proposed Block Diagram

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, Jan-June 2026, Pages: 1810-1817

3.1 Data collection

In a stroke, the brain's blood supply is cut off, which results in the death of brain cells. Hemorrhagic stroke (caused by bleeding) and ischemic stroke (caused by insufficient blood flow) are the two basic forms of stroke. Both may result in abnormal brain function in certain areas. The type of stroke that occurs if symptoms go away in less than a day or two is called a transient ischemic attack (TIA), or mini-stroke. Severe headaches are another complication of hemorrhagic strokes. Stroke symptoms could last for ever. As shown in Figure 2, long-term complications can include pneumonia and bladder control.

id	gender	age	hypertension	heart_disease	ever_married	work_type	Residence_type	avg_glucose_level	bmi	smoking_status	stroke
9046	Male	67	0	1	Yes	Private	Urban	228.69	36.6	formerly smoked	1
51676	Female	61	0	0	Yes	Self-employe	Rural	202.21	N/A	never smoked	1
31112	Male	80	0	1	Yes	Private	Rural	105.92	32.5	never smoked	1
60182	Female	49	0	0	Yes	Private	Urban	171.23	34.4	smokes	1
1665	Female	79	1	0	Yes	Self-employe	Rural	174.12	24	never smoked	1
56669	Male	81	0	0	Yes	Private	Urban	186.21	29	formerly smoked	1
53882	Male	74	1	1	Yes	Private	Rural	70.09	27.4	never smoked	1
10434	Female	69	0	0	No	Private	Urban	94.39	22.8	never smoked	1
27419	Female	59	0	0	Yes	Private	Rural	76.15	N/A	Unknown	1
60491	Female	78	0	0	Yes	Private	Urban	58.57	24.2	Unknown	1
12109	Female	81	1	0	Yes	Private	Rural	80.43	29.7	never smoked	1
12095	Female	61	0	1	Yes	Govt_job	Rural	120.46	36.8	smokes	1
12175	Female	54	0	0	Yes	Private	Urban	104.51	27.3	smokes	1
8213	Male	78	0	1	Yes	Private	Urban	219.84	N/A	Unknown	1
5317	Female	79	0	1	Yes	Private	Urban	214.09	28.2	never smoked	1
58202	Female	50	1	0	Yes	Self-employe	Rural	167.41	30.9	never smoked	1
56112	Male	64	0	1	Yes	Private	Urban	191.61	37.5	smokes	1
34120	Male	75	1	0	Yes	Private	Urban	221.29	25.8	smokes	1
27458	Female	60	0	0	No	Private	Urban	89.22	37.8	never smoked	1

Figure 2: Dataset Information

Qualitative Data

- Find "Female," "Male,"
- Age: The age of the patient
- Hypertension: 1 in the event of its presence and 0 in its absence
- Heart Disease: 1 if the patient has heart problems, 0 otherwise. 5) Everlasting Marriage: "Yes" or "No"
- Type of Job: "Self-Employed," "Govtjov," "Children," "Never Worked," and "Private"
- Which kind of home—"Rural" or "Urban"?
- Average 9) BMI: Body Mass Index Blood Glucose: Average Blood Glucose Level
- "Never Smoked," "Formerly Smoked," "Smokes," or "Unknown" are the available smoking statuses.* Stroke: 1 in case the patient had a stroke, or 0 in the event of a 1.

3.2 Data Preprocessing

Pre-processing using Data Weighted Mean Filtering is a nonlinear noise reduction approach. Take irrelevant values from rows and columns of a learned dataset that has been transformed. Substitutes the median amount value surrounding the data for the value of the central data.

Input: Data weights of feature (Fx)
Output: mean filtered original dataset
Start = type (a);
$Irowweight_{fsum} = colnweight[fw(1)];$
For $(w = 1;\ w < Len(feature\ weight);\ w = w + 1)$
{
$dataweight_{fsum}(w) = colweight[i(fw - 1)] + rowweight[i(1)];$
}
For $(fw = 0;\ w < Len(w);\ fw = fw + 1)$
{
$If\ dataweight_{fsum}(fw) >= dataweight_{f}[len\ (x) - 1/2];$
{
Mean = x [index (I)]
Return
}
}

Where, $w - input\ values\ of\ images,$

Reducing null values is mostly used to improve the corrupted data's characteristics by getting rid of noise. Data Weighted Mean Filtering (DWMF) is a method used to build original datasets by reducing null values.

3.2 Feature selection using multi perception neural network (MPNN)

The technique of extracting features from stroke data is crucial and holds great influence over the outcomes of classification. The required information is taken out of the data and used in the suggested way. High input feature sets can be handled using feature extraction, which also helps to lower dimensionality, identify the most important features various classes.

The stroke data that was taken out of the segmentation data is the input for the Convolutional Layer Similarity Function.

$$O_w \sim \phi(S_{w+1} O_{w+!})$$

$$Sf_1(fw_1, fw_2, \ldots fw_m,) = ||a - f_1\phi\left(f_2\phi(\phi(fw_m))\right)||^2 f + \lambda(f_m)$$

$$w_{x,y} = \begin{cases} \exp\left(-\dfrac{||f_{i,j} - f_{i,j}||^2}{\sigma}\right), if\ f_{i,j} \in \mathrm{CDFS2}(f_{i,j})\ and l_i =\ l_j\ else \\ 0, \end{cases}$$

Here, another feature space is estimated using feature-based spectral input for the subsequent layer. The eigenvalue is an, the regularisation parameter is λ-, and the sigmoid activation function is φ.

MSW MANAGEMENT -**Multidisciplinary, Scientific Work and Management Journal**
ISSN: 1053-7899
Vol. 36 Issue 1, Jan-June 2026, Pages: 1810-1817

$$S_2(f_n) = \sum_{k=1}^{f_n}\sum_{k=1}^{f_n} w_{x,y}\,||s_{1,m} - s_{1,n}||^2 = (S_n^m)$$

S_n^m is the similarity scaled value values, To optimise the convolution dictionary training and classifier for brain stroke data classification.

To maintain the data's behind structural information, the original samples must be rebuilt in the final layer.

$S = \{S1, S2\}$← Pre-process the stroke data
R-{} # relative set
f_1, f_2- Adaptable feature set
S←Choose the stroke information using threshold weights.
Train classifier (values of features, R)
For x=1: feature index no
Value extraction (Stroke data, threshold) ←R ($S_2(f_n)$)
If the feature set for stroke data == threshold values>S
Include feature values to f_1, f_2
End
End for

Retrieving prominent and identifiable characteristics from brain MRI stroke data is essential for precise stroke categorization. The stroke dataset are used to train the model, and several characteristics are integrated to enhance classification performance. Traditional methods are utilised to analyse typical properties in order to balance the support strength levels for each feature; probabilistic models are then employed in order to attain high strength levels for strokes. To ascertain the stroke level of the covered pixel area, a probability measure is derived from the highly scattered stroke data that is collected near the point where T(g) is the original stroke level.

$$T\,(v) = N(v)/\,M$$

$$T = \sum_{v=0}^{L-1} vN(u) = \sum_{n=0}\sum_{n=0}\frac{I(m,k)}{n}$$

$$\sigma_v = \sqrt{\sum_{v=0}^{L-1}(v-\bar{v})^2 h(v)}$$

Support for "u" point or row and column entities with constant equation variance generated from low-contrast stroke data, included in the stroke data model. This enhances the square root change in the stroke data, which is a measure of stroke effectiveness.

3.3 Classification based Gated Convolutional Neural Network (CNN)

The development of medical applications has significantly advanced thanks to the use of gated convolutional neural networks (CNNs), one of the newest technologies in artificial intelligence. Thanks to its superior accuracy and low error rate, CNN began to surpass all other models and gained popularity. This inspired us to forecast the risk of stroke disease using his CNN algorithm. The layers that comprise CNN are fully correlated, pooling, and convolutional layers.

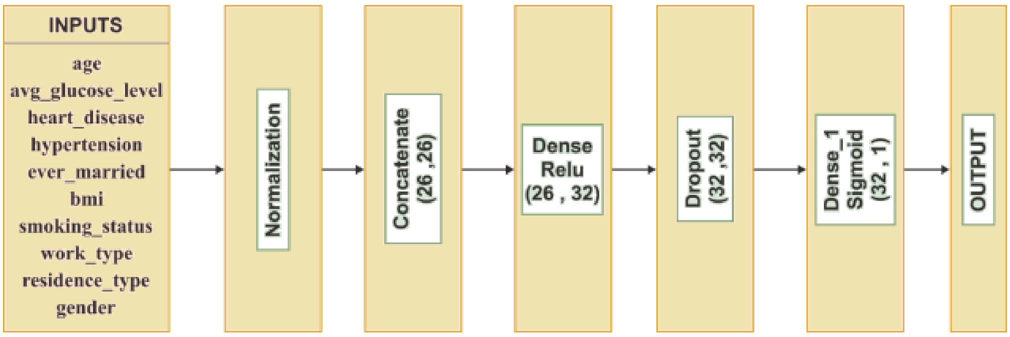

Figure 3: Proposed Algorithm Flow

Our suggested CNN approach uses the model to generate 1D outputs and leverages structured 1D features in the dataset. Ten features from our dataset are used to train the model. The obtained results are sent to a preset dataset to see how the model performs under a specific section or load. Then, using backpropagation to update the learnable bounds, chunks, and loads according to the disaster scenario, the estimates in Figure 3 generally get better. There are unquestionably more learnable boundaries to measure, making it more valuable. The model preparation method necessitates a graphics processing unit (GPU), which raises the computational expense significantly.

Dense points that integrate activation functions for logic to create the shape of gated neural networks.

$R_t = u(a_i x_t + a_h y_{t-1} + a_c z_{t-1} + b_x)$ And

$$S_t = v\left(a_{fx}x_t + a_{hf}h_{t-1} + a_{cf}z_{t-1} + b_f\right)$$

Where is the sigmoid activation function (s) representation? It is z(t-1) that provides the unit cell.An activation unit is any activated neuron zt−1, starting from the sigmoid adaptation S. The formula for $z_{t-1} = f_{wt} \odot z_{t-1} + i_t \odot \tanh(a_{cx}x_t + w_{ch}r_{t-1} + b_c)$.

Where ⊙-multiplication and $tanh\ limit\ (x_t, r_{t-1}) = (a_{cx}x_t + a_{ch}h_{t-1} + b_c)$ in the range (-1, 1).

These attributes are used to compare and obtain the input sequence from a brain stroke data set with an MRI stroke data set. Next, from the output sequence, An activation function from a modified CNN approach with the corresponding threshold weights changed based on the bias value is called a recognition function. The logic processor of the hidden unit is what instructs neurons on how to choose feature weights.

$$r_t = relu(a_{wt}x_t + a_r r_{t-1} + b_r)\ y_t = a_{rc} + b_c$$

The activation function receives the support weight w and the bias value after the CNN function optimises the softmax unit of the bias value.

Neural Network weights matrices

A multiple hidden point threshold, h, supports the building of an adaptive iteration layer.

$$H_t = \left(mod\left(h_{t-1}a_{rr}^{(2)}\right) + mod\left(h_{t-1}a_{rr}^{(1)}\right)\right)/2\ \ldots \quad (40)$$

Gated input layers and activation functions can have active support points in the form of input features.. The x^{th} hidden layer r_t^i.for simplicity $a_{rr}^{(1)}$ *and* $a_{rr}^{(2)}$ for a and b replacement.

$$h_t^i = \left(mod\ (\sum jh_{t-1}^j U_{ji}) + \quad mod\ (\sum jh_{t-1}^j V_{ji})\right)/2\ \ldots \qquad (41)$$

With t as the The time operating unit for a layer of k hidden neurons, t-1 at the posterior feature limit can be calculated from (i) the first derivative of the hidden neurons. As a result, the recurrent neural network becomes more adept at differentiating between immature and mature inference as well as threshold values for data classification. Continuous feature support weights logically drive the threshold mean points for the segmentation classes. In order to recognize continuous edges and categorize strokes, Gated generates neuronal states in every layer.

4. RESULT AND DISCUSSION

The proposed model is built with TensorFlow, whereas Keras is written in Python. A UCI repository is made using a Jupiter notebook containing an existing dataset. Based on stroke data, a unique system is provided to determine disease or no disease.

Table 1: Proposed Simulation Parameters

Simulation Parameters	Proposed Values
Dataset Name	Brain stroke dataset
Tool Used	Anaconda
Language	Python
No. of data	1500
Trained data	1000
Ttest data	500

Brain stroke dataset processed to evaluate the efficacy of the suggested approach is described in Table 1. The amount of test and training stroke data as well as the number of strokes assess the stroke detection categorization.

$$Sensitivity\ (S) = TP\ /\ (TP + FP) * 100$$

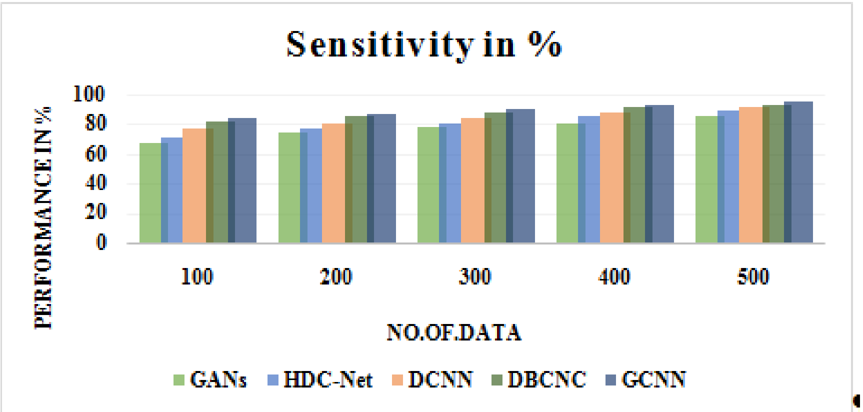

Figure 4: Sensitivity Performance

The proposed implementation works better than previous algorithms, as seen in Figure 4's comparison of true positive accuracy sensitivity values for various approaches. Compared to earlier techniques (XGBoost, LSTM, CNN, and SVM approaches), the suggested CNN method is less sensitive to stroke data even though it reaches a sensitivity performance of 92%. On the other hand, the sensitivity score of 92% for the suggested CNN method is higher than that of earlier approaches.

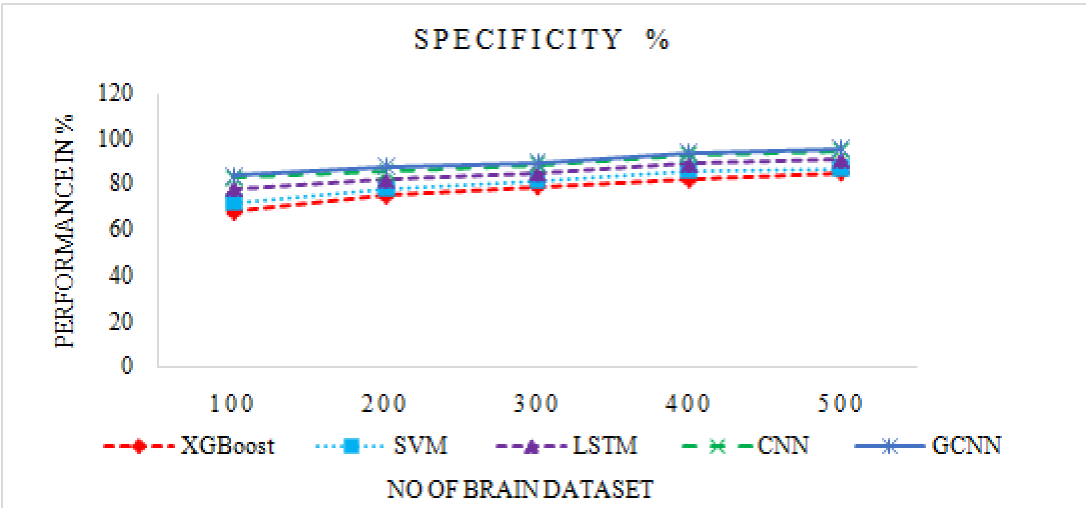

Figure 5: Specificity Performance

When compared to alternative approaches, Figure 5 displays an accurate study of specificity and true negative accuracy, demonstrating the good performance of the suggested implementation. While earlier techniques (such as XGBoost, LSTM, CNN, and SVM approaches) had lower specificity for stroke data, the suggested CNN method produced a sensitivity performance of 93%; however, the suggested method A CNN has a high sensitivity of 93%, which is superior toprevious techniques.

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, Jan-June 2026, Pages: 1810-1817

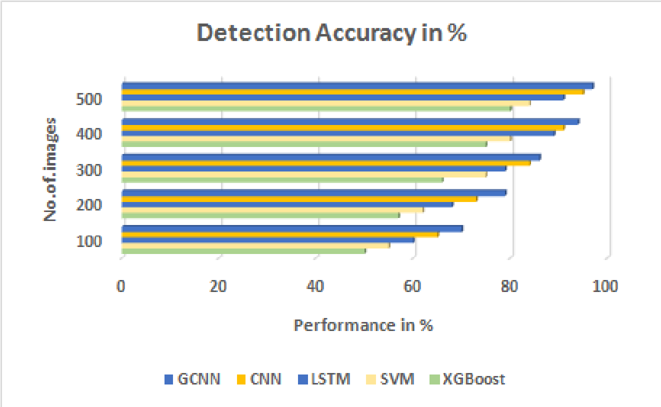

Figure 6: Performance of Detection Accuracy

A comparison of the detection accuracy values of various strategies is shown in Figure 6, where the suggested implementation performs better than alternative algorithms. The suggested method, CNN, has a detection accuracy of up to 92%, which is higher than earlier approaches.

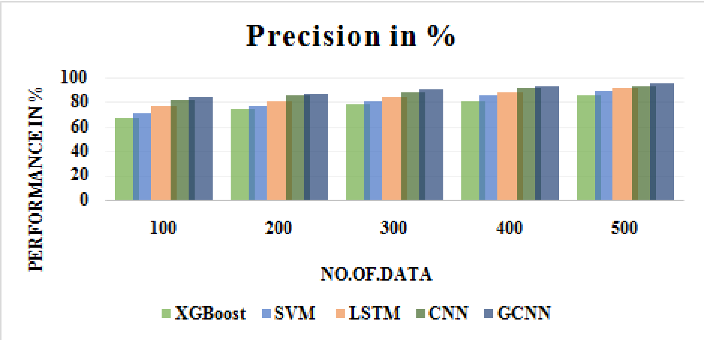

Figure 7: Analysis of Precision Performance

The suggested implementation performs better than other algorithms when comparing the accuracy values for genuine positive accuracy of various approaches, as shown in Figure 7. In contrast to other techniques (XGBoost, LSTM, CNN, and SVM approaches), the suggested CNN method achieved a sensitivity performance of 92% on stroke data. However, compared to earlier approaches, the suggested method A CNN has a sensitivity of up to 95%.

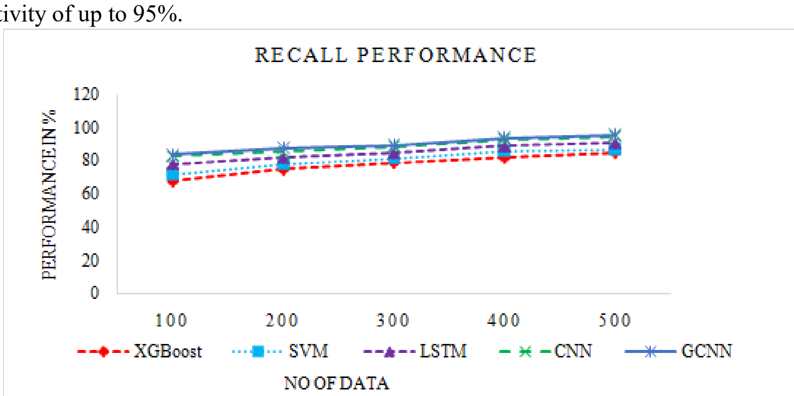

Figure 8: Analysis of Recall Performance

The suggested implementation performs better in Figure 8, which displays the repeatability analysis of true negative accuracy in comparison to alternative approaches. While earlier techniques (such as XGBoost, LSTM, CNN, and SVM approaches) had lower specificity for stroke data, the suggested CNN method produced a sensitivity performance of 92%; however, the suggested method A CNN has a high sensitivity of 92%, which is greater than previous techniques.

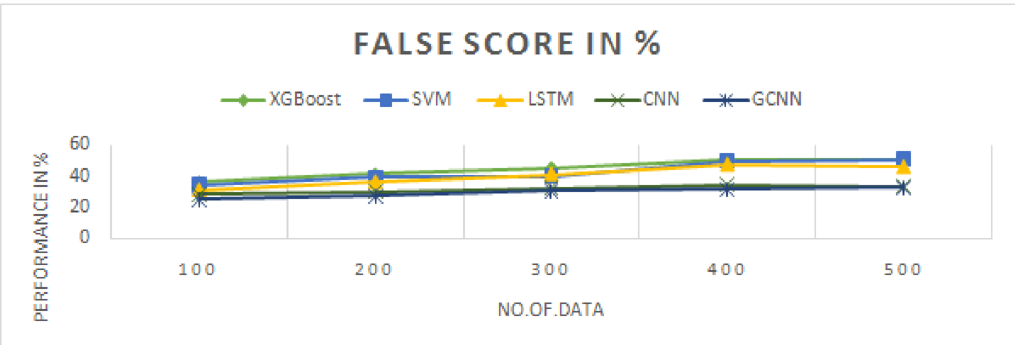

Figure 9: Analysis of False Score

The suggested implementation's error rate performance is low when compared to other algorithms, as can be seen by comparing the error rate numbers of the various approaches displayed in Figure 9. The suggested CNN system reduces mistake rate by 30.5 %.

5. CONCLUSION

In this study, stroke data is used to classify and segment the risk factors. Classification kinds using machine learning models. For segmentation, autoencoder-decoder and CNN are utilised. 500 data points were used to evaluate the model in this study, and the data set was in.CSV format. The categorization model had a 92% accuracy rate. The examines' results show how crucial these deep neural networks are for stroke diagnosis. Thus, compared to the conventional approach, the recommended method produces results that are more accurate. Although machine learning has great promise for the field of medical imaging, there are numerous obstacles to overcome. Constructed using extensive medical datasets to increase the precision of segmentation algorithms and employ many machine learning sophisticated pre-trained models to determine the precise type of stroke.

Acknowledgement: Nil
Conflicts of Interest: The authors report there are no competing interests to declare.
Funding Sources: No funding was received to assist with the preparation of this manuscript.

REFERENCE

[1] M. Lee, J. -H. Jeong, Y. -H. Kim and S. -W. Lee, "Decoding finger tapping with the affected hand in chronic stroke patients during motor imagery and execution", IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 29, pp. 1099-1109, 2021.
[2] N. Rungsirisilp and Y. Wongsawat, "Applying combined action observation and motor imagery to enhance classification performance in a brain computer interface system for stroke patients", IEEE Access, vol. 10, pp. 73145-73155, 2022.
[3] N. Vivaldi, M. Caiola, K. Solarana and M. Ye, "Evaluating performance of EEG data driven machine learning for traumatic brain injury classification", IEEE Transactions on Biomedical Engineering, vol. 68, no. 11, pp. 3205-3216, 2021.
[4] J. Leng et al., "Time frequency space EEG decoding model based on dense graph convolutional network for stroke", IEEE Journal of Biomedical and Health Informatics, vol. 28, no. 9, pp. 5214-5226, 2024.
[5] S. Zhang et al., "Learning EEG representations with weighted convolutional siamese network: a large multi-session post-stroke rehabilitation study", IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 30, pp. 2824-2833, 2022.
[6] M. Lee, H. -Y. Park, W. Park, K. -T. Kim, Y. -H. Kim and J. -H. Jeong, "Multi-Task heterogeneous ensemble learning-based cross subject EEG classification under stroke patients", IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 32, pp. 1767-1778, 2024.
[7] M. A. Saleem et al., "Innovations in stroke identification: A machine learning-based diagnostic model using neuroimages", IEEE Access, vol. 12, pp. 35754-35764, 2024.
[8] S. Datta, C. K. Karmakar, B. Yan and M. Palaniswami, "Novel measures of similarity and asymmetry in upper limb activities for identifying hemiparetic severity in stroke survivors", IEEE Journal of Biomedical and Health Informatics, vol. 25, no. 6, pp. 1964-1974, 2021.
[9] G. Zhu, A. Bialkowski, L. Guo, B. Mohammed and A. Abbosh, "Stroke classification in simulated electromagnetic imaging using graph approaches", IEEE Journal of Electromagnetics, RF and Microwaves in Medicine and Biology, vol. 5, no. 1, pp. 46-53, 2021.
[10] S. Sakri et al., "An Improved concatenation of deep learning models for predicting and interpreting ischemic stroke", IEEE Access, vol. 12, pp. 53189-53204, 2024.
[11] Sirsat, M. S., Fermé, E., and Camara, J, "Machine learning for brain stroke: a review", Journal of stroke and cerebrovascular diseases, 2020.
[12] Sailasya, G., & Kumari, G. L. A, "Analyzing the performance of stroke prediction using ML classification algorithms", International Journal of Advanced Computer Science and Applications, 2021.
[13] Asit Subudhi, Manasa Dash, Sukanta Sabut, "Automated segmentation and classification of brain stroke using expectation-maximization and random forest classifier, Biocybernetics and Biomedical Engineering, Volume 40, Issue 1, 2020.
[14] Pratama, M. R., Suryana, A. L., Alfiansyah, G., Olivia, Z., Nurmawati, I., & Destarianto, P, "Diagnosis of stroke and diabetes mellitus with classification techniques using decision tree method", International Journal of Health and Information System, 2(1), 1-8, 2024.
[15] Abbaoui, W., Retal, S., Ziti, S., El Bhiri, B., & Moussif, H, "Ischemic Stroke Classification Using VGG-16 Convolutional Neural Networks: A Study on Moroccan MRI Scans", International Journal of Online and Biomedical Engineering, 20(2).
[16] Jarapala Parvathi. 2024. "Machine Learning Based Brain Stroke Prediction Using Light Gradient Boosting Machine Algorithm," International Journal of Intelligent Systems and Applications in Engineering 12 (4):2670 -. https://ijisae.org/index.php/IJISAE/article/view/6734.
[17] Senjuti Rahman, and Ajay Krishno Sarkar, 2023, "Prediction of Brain Stroke using Machine Learning Algorithms and Deep Neural Network Techniques", pp.23-30.
[18] Ferdous, M. J., & Shahriyar, R. (2024). An ensemble convolutional neural network model for brain stroke prediction using brain computed tomography images. *Healthcare Analytics*, *6*, 100368. https://doi.org/10.1016/j.health.2024.100368.
[19] Choi, A., Park, J., Jun, A., Pyo, S., Cho, H., Lee, S., & Yu, H. (2021). Deep Learning-Based Stroke Disease Prediction System Using Real-Time Bio Signals. *Sensors (Basel, Switzerland)*, *21*(13), 4269. https://doi.org/10.3390/s21134269.
[20] Srinivas, A., & Mosiganti, J. P. (2023). A brain stroke detection model using soft voting-based ensemble machine learning classifier. *Measurement: Sensors*, *29*, 100871. https://doi.org/10.1016/j.measen.2023.100871
[21] Saleem, Muhammad Asim, et al. "Innovations in stroke identification: A machine learning-based diagnostic model using neuroimages." IEEE Access (2024).
[22] O Ramya Teja, B. Rajya Laxmi, P. Hema Sree, J. Madhurika, and R. Bhavani, "DEEP LEARNING BASED BRAIN STROKE PREDICTION," JAC A Journal of Composition Theory, Volume XVI, Issue-7, JULY 2023, ISSN: 0731-6755.
[23] Hung CY, Lin CH, Lan TH, Peng GS, Lee CC. Development of an intelligent decision support system for ischemic stroke risk assessment in a population-based electronic health record database. PLoS One. 2019 Mar 13;14(3): e0213007. doi: 10.1371/journal.pone.0213007. PMID: 30865675; PMCID: PMC6415884.
[24] Fang, G., Huang, Z., & Wang, Z. (2022). Predicting Ischemic Stroke Outcome Using Deep Learning Approaches. *Frontiers in Genetics*, *12*, 827522. https://doi.org/10.3389/fgene.2021.827522
[25] Dhillon S, Bansal C, Sidhu B. Machine Learning Based Approach Using XGboost for Heart Stroke Prediction. in International Conference on Emerging Technologies: AI, IoT, and CPS for Science & Technology Applications, September 06–07, 2021.