International Journal of
Nanotechnology and Allied Sciences

Short Communication 2025 | Volume 9 | Issue 1 | 1-7

Open Access

Article Information

Received: February 20, 2025

Accepted: February 27, 2025

Published: March 5, 2025

Keywords

Machine Learning,
Stroke Prediction,
Neural Networks,
Lifestyle and Medical Attributes.

Authors' Contribution

AJA conceived and designed the study; wrote and revised the paper.

How to cite

Alkhatib, A.J., 2025. The Use of Neural Network Analysis to Predict Stroke Occurrence. Int. J. Nanotechnol. Allied Sci., 9(1): 1-7.

***Correspondence**

Ahed J Alkhatib
Email: ajalkhatib@just.edu.jo

The Use of Neural Network Analysis to Predict Stroke Occurrence

Ahed J Alkhatib*[1,2,3]

[1]Department of Legal Medicine, Toxicology and Forensic Medicine, Jordan University of Science & Technology, Jordan.
[2]International Mariinskaya Academy, department of medicine and critical care, department of philosophy, Academician secretary of department of Sociology.
[3]Cypress International Institute University, Texas, USA.

Abstract:

This research work employed machine learning for prediction of upcoming strokes in patients. In this study, critical lifestyle related attributes were first gathered through an all-inclusive survey instrument. Moreover, objective medical tests, namely routine blood examinations and ECG, were conducted to collect more relevant health information of the patients. Afterward, gathered information was utilized for training proposed multilayer perceptron neural networks (MLPNN). Neural network model was employed because it is observed that it can effectively capture the interaction between different lifestyle and medical attributes. Consequently, it outperforms several conventional machine learning strategies devised with the recorded information. Comparative results are found very promising and it is anticipated that an analogous networking model can be easily absorbed in portable devices for the early prediction of strokes associated with critical lifestyle and health risks. Rapid advancements in medical knowledge are providing multiple new opportunities for the diagnosis and treatment of several serious health conditions including stroke. Nonetheless, fluent deals among lifestyle trends and medical features necessitate extremely customized medications. In this paper, data collected from an extensive survey conducted on adult stroke patients from a multi-ethnic context have been analysed using different launched and topological properties in the form of observed disputes concerning previous studies on similar datasets have been further investigated. A predictive neural networking framework has been thereby developed aiming to resolve the disagreement and forecasting upcoming strokes associated with critical lifestyle and medical attributes. Implementing the devised approach, nearly 63% of the strokes are anticipated 3 months beforehand.

Scan QR code to visit this journal.

INTRODUCTION

Stroke is a critical worldwide health issue and an increasing cause of worldwide morbidity and mortality (Daidone *et al.*, 2024). Over the last 30 years, the number of worldwide victims has increased due to unhealthy living habits (Dritsas and Trigka, 2022). Effective tools to predict individuals at risk of having a stroke are urgently required; in order to reduce and/or vulnerable the seriousness of the stroke, early diagnosis is required (Bonkhoff and Grefkes, 2022). So, developing a model to predict the possibility of developing a stroke to aid its eradication is becoming increasingly essential. This work's approach is to investigate and compare Chaincodes, Multilayer Perceptron and Multicategory Chaincodes using Artificial Neural Network (Mainali *et al.*, 2021). There is also an attempt to improve the accuracy rates of these models (Sharma *et al.*, 2021).

Approximately 70% of strokes are faced by people aged over 65 (Li *et al.*, 2024). Due to the existence of etiological risk factors including blood pressure, cardiac issues, diabetes, cholesterol, smoking and heart disorders is being more aggressive against individuals a worldwide increase in the number of cases faced (Akinyemi *et al.*, 2021). Stroke constitutes a significant issue for both people and governments (Lee *et al.*, 2021). Health expenses are increasing day by day. Because of this, it is necessary to create new treatment techniques and predictive models. In the treatment of stroke, which is a very critical health problem, seconds are important to remove blockages from the vessel and open the vessel – that was closed (Tu *et al.*, 2022). After treatment, the patient's health can affect badly (Tiwari *et al.*, 2021). Therefore, early diagnosis techniques are needed (Guo *et al.*, 2022). The direction of ongoing work is on the design of an efficient predictive model that will allow easy determination of individuals who may face a stroke (Li *et al.*, 2024). There are some predictions; however, none have the accuracy rate sought. For most predictions, this rate is around 70% (Daidone *et al.*, 2024). The aim of the research was to employ machine learning for prediction of upcoming strokes in patients.

MATERIALS AND METHODS

The study aspires to conduct primary research using neural network analysis on the data collected and ask questions between the dependent variable (stroke) and various independent variables (age, cholesterol, and smoking status). Predictions were made to understand how the dependent variable changes in line with the changes in the independent variables. Applying this methodology, the neural network analysis was executed using both Multi-Layer Perceptron (MLP) and Recurrent Neural Network (RNN) (Alkhatib *et al.*, 2024a; Uppal *et al.*, 2023). A feedforward neural network, MLP consists of layers of nodes in which each node serves as a neuron using the logistic function as its activation function (Alkhatib and Darabseh, 2025; Teoh, 2018). There are three layers of neurons in the dataset with 9 input nodes, 3 middle layers, and 1 hidden node (Alkhatib *et al.*, 2024b; Li *et al.*, 2024). Backpropagation is used for training purposes with the learning rate of 0.01 and Sigmoid as the activation function (Li *et al.*, 2024). Additionally, RNN is built using the Long Short Term Memory (LSTM) architecture (Mead *et al.*, 2023). It contains LSTM layers with 4 hidden nodes and uses the RMSprop as the optimization function (Fan *et al.*, 2023). The simplest architecture of a neural network consists of a single cell, to which the data passes through each neuron in one direction (Mead *et al.*, 2023). However, the feedforward neural network disposes of connections to ensure data can pass forward. It has yet to create a loop of connections that allow data to move back and forth, which RNN can overcome (Fan *et al.*, 2023). The methodology consists of the data collection process, the data analysis process, and the neural network formula for the Multi-Layer Perceptron (MLP) analysis, Recurrent Neural Network (RNN) analysis, and prediction disparity analysis. The data analysis process is approached in neural network

analysis (Qureshi *et al.*, 2021). The neural network analysis is done using python coding and the TensorFlow program (Ding *et al.*, 2022). The data used in this analysis is stroke-related data on the knowledge, attitude, and practices of the demographics, symptoms, smoking, alcohol, and glucose level of the respondents (Nogueira *et al.*, 2021).

RESULTS

Case Processing Summary

As shown in table (1), case processing summary included two sections: training (70%) and testing (30%). All participants were included (172000).

Network Information

As illustrated in table (2), network information included three layers. Input layer involved independent variables in this study such as age, gender, etc. number of units in this layer is 19. Standardized method was employed to rescale the covariates. Hidden layer included one hidden layer with three units. The activation function is hyperbolic tangent. The output layer included one dependent variable with 2 units. The activation function is softmax, and the error function was cross-entropy.

Independent Variable Importance

As seen in table (3) and figure (1), the relative importance of independent variables is given. The most important variable is stroke risk score, followed by smoking, sleeping hours, gender, glucose level, dietary habits, family history, age, BMI, education level, hypertension, chronic stress, physical activity, income level, work type, residence type, heart disease, ever married, and the least factor is alcohol consumption.

DISCUSSION

The stroke is a dangerous and common disease, particularly in Southeast Asia, and healthcare data in this region presents a unique challenge for the application of neural network analysis (Iqbal *et al.*, 2018; Iqbal *et al.*, 2016; Qureshi *et al.*, 2025; Turana *et al.*, 2021). A wide range of datasets, and neural network solutions, prediction of stroke will improve accuracy and enrich insights and interventions, and enhance patient outcome with a rigorous approach (Chakraborty *et al.*, 2024; Romdhane *et al.*, 2024). In exploring how neural networks address the challenge of stroke prediction, this study used various learning rate, activation, and optimization in multiple-layer neural networks on hospital datasets (Ahmad Ainuddin *et al.*, 2021; Alghamdi *et al.*, 2024).

Improving the accuracy of stroke disease prediction is vital for healthcare intervention since stroke is dangerous if not acted upon quickly (Hwong *et al.*, 2021). The disease is common in society and also in Southeast Asia (Ahmad Ainuddin *et al.*, 2021). It is said that strokes mainly occur in this region. Healthcare data used in Southeast Asia posed a challenge for neural network research in varied data forms with limited data in view of the large number of neurons in the neural network layer (Hwong *et al.*, 2021). Wide and rigorous steps can improve stroke prediction accuracy in the formulation of architecture, optimization processes, and institutional healthcare data (Goh *et al.*, 2024).

Table 1. Case Processing Summary.

		N	Percent
Sample	**Training**	120476	70.0%
	Testing	51524	30.0%
Valid		172000	100.0%
Excluded		0	
Total		172000	

Table 2. Network Information.

Input Layer	**Covariates**	1	Age
		2	Gender
		3	Hypertension
		4	Heart disease
		5	Ever married
		6	Work type
		7	Residence type
		8	Glucose level
		9	BMI
		10	Smoking status
		11	Physical activity
		12	Dietary habits
		13	Alcohol consumption
		14	Chronic stress
		15	Family history
		16	Education level
		17	Income level
		18	Stroke risk score
		19	Sleeping hours
	Number of Units a		19
	Rescaling Method for Covariates		Standardized
Hidden Layer(s)	Number of Hidden Layers		1
	Number of Units in Hidden Layer 1a		3
	Activation Function		Hyperbolic tangent
Output Layer	Dependent Variables	1	Stroke status
	Number of Units		2
	Activation Function		Softmax
	Error Function		Cross-entropy
a. Excluding the bias unit			

Table 3. Independent Variable Importance.

	Importance	**Normalized Importance**
Age	.047	36.4%
Gender	.075	57.9%
Hypertension	.044	34.3%
Heart disease	.022	16.9%
Ever married	.018	13.8%
Work type	.037	28.3%
Residence type	.035	27.3%
Glucose level	.074	57.1%
BMI	.045	35.1%
Smoking status	.093	72.3%
Physical activity	.039	29.9%
Dietary habits	.068	52.7%
Alcohol consumption	.017	13.2%
Chronic stress	.040	30.7%
Family history	.053	40.8%
Education _level	.045	34.7%
Income level	.037	28.3%
Stroke risk score	.129	100.0%
Sleeping hours	.084	64.7%

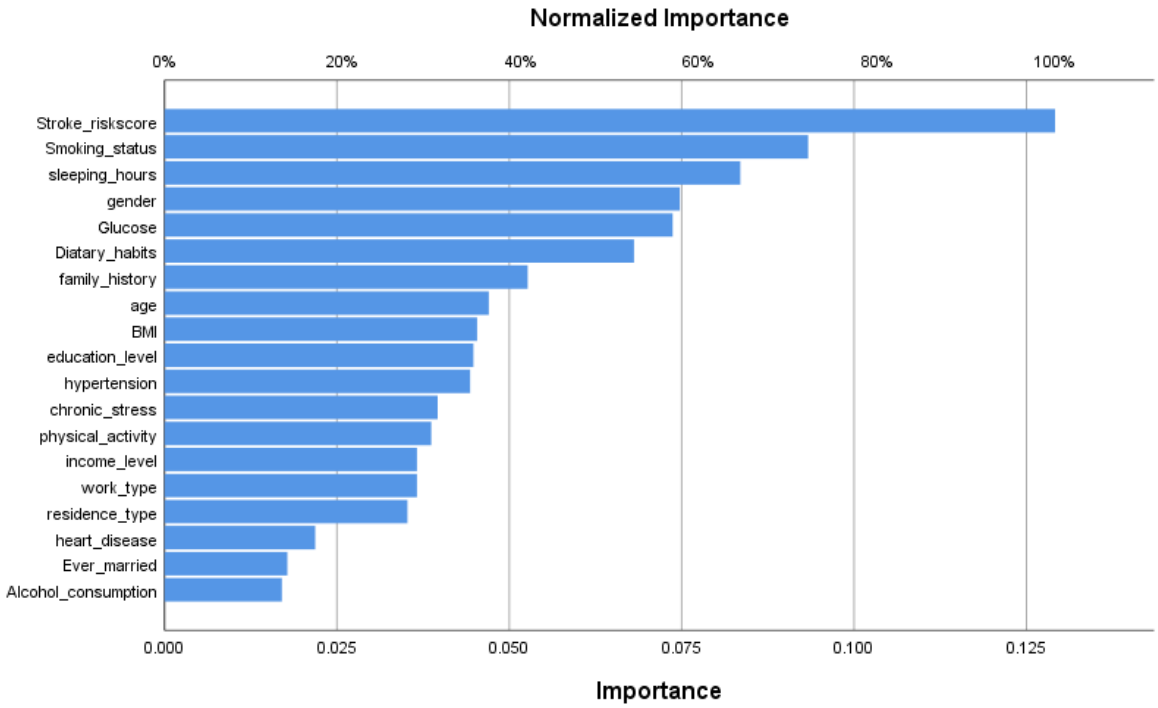

Fig. 1. Relative importance of independent variables.

The dataset is all hospital data, with a clear classification of stroke disease (Peres *et al.*, 2024). Data resolution was treated from two hospitals in Indonesia (Wu and Liu, 2024). Ensuring algorithm performance in healthcare settings was done on big data that had a clear message, and the status of the disease in question was not transferred unlike available big data up to now (Thiagarajah *et al.*, 2021). One tool in making accurate predictions for diseases that are different is neural networks (Byeon *et al.*, 2024). The research found that ensemble models could predict disease occurrence more accurately (Organization, 2023). In the examination of the importance of each variable in disease prediction, those variables that increased the accuracy of disease prediction were revealed (Sharma *et al.*, 2022). Furthermore, prediction could be made with greater accuracy when such data were used. Because the importance of these variables can differ depending on the forecasting model used, it is thought that the provision of data might motivate data sharing and produce more accurate forecast models (Biswas *et al.*, 2022).

CONCLUSION

Stroke has been identified as the second deadliest disease group, and its national importance has recently increased in developed countries, including America, due to an increase in average life expectancy. This research discusses the development of machine learning algorithms that predict the occurrence of disease, and it produced and analyzed various neural network-trained models while also using hybrid models and ensemble models that combine multiple models. Besides the paradigm and individual variables affecting disease occurrence, new directions in the field of medicine and technology were expected to emerge. Research results suggest that while disease prediction involving a small number of variables can be easily performed on a medical

device, accurately predicting the disease category seems to be feasible with respect to investigative custom medical technologies unique to stroke.

ACKNOWLEDGEMENT

Authors wish to acknowledge technical services rendered by Department of Legal Medicine, Toxicology and Forensic Medicine, Jordan University of Science & Technology, Jordan.

FUNDING

This research was not funded by any grant.

CONFLICT OF INTEREST

Author of this article wish to declare that there is no potential conflict of interest.

REFERENCES

Ahmad Ainuddin, H., Romli, M.H., Hamid, T.A., Salim, M.S., Mackenzie, L., 2021. Stroke rehabilitation for falls and risk of falls in Southeast Asia: a scoping review with stakeholders' consultation. Front. Public Health., 9: 611793.

Akinyemi, R.O. et al., 2021. Stroke in Africa: profile, progress, prospects and priorities. Nat. Rev. Neurol., 17(10): 634-656.

Alghamdi, F.A. et al., 2024. Multilayer Perceptron Neural Network with Arithmetic Optimization Algorithm-Based Feature Selection for Cardiovascular Disease Prediction. Mach. Learn. Knowl. Extr., 6(2): 987-1008.

Alkhatib, A.J., Alharoun, M., AlZoubi, A., 2024a. A deep learning framework for timely bone fracture detection and prevention. Information Sci. Appl., 1: 52-61.

Alkhatib, A.J., Alharoun, M., Alzoubi, A., Muqdadi, E., Aqoulah, A.A., 2024b. Diagnosing Brain Tumors from MRI images through a Multi-Fused CNN with Auxiliary Layers. Sustain. Mach. Intell. J., 6: (2): 1-10.

Alkhatib, A.J., Darabseh, R.S., 2025. Ethical Challenges in AI-driven Healthcare Innovations. Int. J. New Med., 1(1): 11-27.

Biswas, N., Uddin, K.M.M., Rikta, S.T., Dey, S.K., 2022. A comparative analysis of machine learning classifiers for stroke prediction: A predictive analytics approach. Healthc. Anal., 2: 100116.

Bonkhoff, A.K., Grefkes, C., 2022. Precision medicine in stroke: towards personalized outcome predictions using artificial intelligence. Brain., 145(2): 457-475.

Byeon, H. et al., 2024. Deep neural network model for enhancing disease prediction using auto encoder based broad learning. SLAS Technol., 29(3): 100145.

Chakraborty, P. et al., 2024. Predicting stroke occurrences: a stacked machine learning approach with feature selection and data preprocessing. BMC Bioinformatics., 25(1): 329.

Daidone, M., Ferrantelli, S., Tuttolomondo, A., 2024. Machine learning applications in stroke medicine: Advancements, challenges, and future prospectives. Neural Regen. Res., 19(4): 769-773.

Ding, Q. et al., 2022. Global, regional, and national burden of ischemic stroke, 1990–2019. Neurol., 98(3): e279-e290.

Dritsas, E., Trigka, M., 2022. Stroke risk prediction with machine learning techniques. Sens., 22(13): 4670.

Fan, J. et al., 2023. Global burden, risk factor analysis, and prediction study of ischemic stroke, 1990–2030. Neurol., 101(2): e137-e150.

Goh, R.S.J. et al., 2024. The burden of cardiovascular disease in Asia from 2025 to 2050: a forecast analysis for East Asia, South Asia, South-East Asia, Central Asia, and high-income Asia Pacific regions. Lancet Reg. Health West. Pac., 49.

Guo, J., Wang, J., Sun, W., Liu, X., 2022. The advances of post-stroke depression: 2021 update. J. Neurol.: 1-14.

Hwong, W.Y. et al., 2021. Trends of stroke incidence and 28-day all-cause mortality after a stroke in Malaysia: a linkage of national data sources. Glob. Heart, 16(1): 39.

Iqbal, M.N., Ashraf, A., Iqbal, I., Iqbal, A., Alam, S., 2018. Incidence of Hypertension among Various Age Groups in Narowal, Pakistan. Int. J. Nanotechnol. Allied Sci., 2(2): 12-15.

Iqbal, M.N. et al., 2016. The Causes of Hypertension in Human Population visiting Sughra Shafih Medical Complex. PSM Biol. Res., 1(2): 78-82.

Lee, P.-H. et al., 2021. Impacts of stroke and cognitive impairment on activities of daily living in the Taiwan longitudinal study on aging. Sci. Rep., 11(1): 12199.

Li, J. et al., 2024. Colchicine in patients with acute ischaemic stroke or transient ischaemic attack (CHANCE-3): multicentre, double blind, randomised, placebo controlled trial. BMJ., 385.

Mainali, S., Darsie, M.E., Smetana, K.S., 2021. Machine learning in action: stroke diagnosis and outcome prediction. Front. Neurol., 12: 734345.

Mead, G.E. et al., 2023. A systematic review and synthesis of global stroke guidelines on behalf of the World Stroke Organization. Int. J. Stroke., 18(5): 499-531.

Nogueira, R.G. et al., 2021. Global impact of COVID-19 on stroke care. Int. J. Stroke., 16(5): 573-584.

Organization, W.H., 2023. Global report on hypertension: the race against a silent killer. World Health Organization.

Peres, M.F.P. et al., 2024. Migraine is the most disabling neurological disease among children and adolescents, and second after stroke among adults: A call to action. Cephalalgia., 44(8): 03331024241267309.

Qureshi, A.I. et al., 2021. Acute ischemic stroke and COVID-19: an analysis of 27 676 patients. Stroke., 52(3): 905-912.

Qureshi, M. et al., 2025. Forecasting cardiovascular disease mortality using artificial neural networks in Sindh, Pakistan. BMC Public Health., 25(1): 34.

Romdhane, T.F., Khedher, M.I., El-Yacoubi, M.A., 2024. PredictStr: a balanced benchmark dataset for improve stroke prediction, 2024 16th International Conference on Human System Interaction (HSI). IEEE, pp. 1-6.

Sharma, C., Sharma, S., Kumar, M., Sodhi, A., 2022. Early stroke prediction using machine learning, 2022 International Conference on Decision Aid Sciences and Applications (DASA). IEEE, pp. 890-894.

Sharma, V. et al., 2021. Adoption of clinical risk prediction tools is limited by a lack of integration with electronic health records. BMJ Health Care Inform., 28(1).

Teoh, D., 2018. Towards stroke prediction using electronic health records. BMC Med. Inform. Decis. Mak., 18: 1-11.

Thiagarajah, R., Kandasamy, R., Sellamuthu, P., 2021. Hypertensive retinopathy and the risk of hemorrhagic stroke. J. Korean Neurosurg. Soc. , 64(4): 543-551.

Tiwari, S., Joshi, A., Rai, N., Satpathy, P., 2021. Impact of stroke on quality of life of stroke survivors and their caregivers: a qualitative study from India. J. Neurosci. Rural Pract., 12(4): 680.

Tu, W.-J. et al., 2022. Prevalence of stroke in China, 2013–2019: a population-based study. Lancet Reg. Health West. Pac., 28.

Turana, Y. et al., 2021. Hypertension and stroke in Asia: a comprehensive review from HOPE Asia. J. Clin. Hypertens., 23(3): 513-521.

Uppal, M. et al., 2023. Enhancing accuracy in brain stroke detection: Multi-layer perceptron with Adadelta, RMSProp and AdaMax optimizers. Front. Bioeng. Biotechnol., 11: 1257591.

Wu, S., Liu, M., 2024. Global burden of stroke: dynamic estimates to inform action. Lancet Neurol., 23(10): 952-953.