EAI Endorsed Transactions
on Internet of Things
Research Article EAI.EU

# Machine Learning Based Stroke Predictor Application

R. Kishore Kanna*[1], Ch. Venkata Rami Reddy[2], Bhawani Sankar Panigrahi[3], Naliniprava Behera[4], Sarita Mohanty[5]

[1]Department of Biomedical Engineering, Jerusalem College of Engineering (Autonomous), Chennai, Tamil Nadu
[2]School of Computer Science & Engineering, VIT-AP University, Amaravati, India
[3]Department of Information Technology, Vardhaman College of Engineering (Autonomous), Hyderabad, Telangana, India,
[4]School of Computer Engineering, KIIT Deemed to be University, Bhubaneswar, Odisha, India
[5]Department of Computer Application, CPGS, OUAT Bhubaneswar, Odisha, India

## Abstract

When blood flow to the brain stops or slows down, brain cells die because they don't get enough oxygen and nutrients. This condition is known as an ischemic stroke. It is now the biggest cause of death in the whole planet. Examining the afflicted people has shown a number of risk variables that are thought to be connected to the stroke's origin. Numerous studies have been conducted to predict the illnesses associated with stroke using these risk variables. The prompt identification of various warning symptoms associated with stroke has the potential to mitigate the severity of the stroke. The utilization of machine learning techniques yields prompt and precise predictive outcomes. Although its uses in healthcare are expanding, certain research domains have a stronger need for more study. We think that machine learning algorithms may aid in a deeper comprehension of illnesses and make an excellent healthcare partner. The textual dataset of numerous patients, which includes many medical variables, is gathered for this study. The missing values in the dataset are located and dealt with during processing. The dataset is used to train machine learning algorithms including Random Forest, Decision Tree classifier, and SVM. The method that delivers the greatest accuracy for our dataset is then selected once the accuracy of the algorithms has been determined. This aids patients in determining the likelihood of a brain stroke and ensuring they get the right medical attention.

**Keywords:** Stroke, SVM, Healthcare, Machine Learning, ML

Received on 12 December 2023, accepted on 07 March 2024, published on 12 March 2024



doi: 10.4108/eetiot.5384

*Corresponding author. Email: kishorekanna007@gmail.com

## 1. Introduction

The patient, their social network, their family, and their place of employment are all impacted by stroke. In India, it is the one of the leading reasons of death. Because of changes in medical science, it may now be possible to anticipate a stroke using machine learning [1]. Machine learning algorithms are helpful for assessing data and producing precise predictions [2]. The majority of earlier stroke research concentrated on foretelling heart attacks. Brain stroke has gotten little to no attention [3]. Machine learning algorithms are useful for providing accurate analysis and making accurate predictions [4]. Heart rate prediction has been a major focus of previous study on the subject of stroke [5]. Stroke research hasn't been studied much. The idea behind the study is to use machine learning to anticipate the incidence of strokes [6].

Machine learning is one of the effective testing technologies founded on training and testing [7]. AI includes machine learning as one of its subsets, a broad learning discipline in which automata imitate human abilities. This combination of technologies is referred to as machine intelligence [8]. On the other hand, machine learning systems are trained to comprehend data analysis and utilization [9].

According to the concept of machine learning, which learns from natural phenomena and objects, the biological parameters used in this project as testing data include hypertension, heart disease, sex, age, and others [10]. A dataset containing a range of physiological traits has been chosen as. These characteristics will be examined subsequently and utilized to make concluding predictions [11].

For the purpose of understanding the machine learning model, the dataset is first cleaned up. Data preparation is the procedure at this point. The zero values in the data record are verified and input to

accomplish this. The dataset is divided into training and test data after pre-processing [12].

Then, a model is created utilizing this fresh data and a variety of classification techniques [13]. To choose the best trained model for the prediction, the accuracy of each of these techniques is computed and compared. This model's weakness is that it was developed using material information rather than a temporal brain image [14].

## 2. Related Works

In the Stroke MD stroke grouping and prediction system, the study talks about how to make an application interface for neurologists to use to manage and see related medical data [15]. The goal of the system is to make it easier to use a forecasting model to get information and add visible data. The precision of the exhibited data, the arrangement of the scrutinized data (including risk factor groups, clustered groups, data alerts, conclusions, etc.), and the efficacy of user software interaction were significant factors that influenced system design determinations, specifically the design of the user interfaces for doctors and patients, input modalities, and user software interaction techniques. These factors ultimately led to the development of a hybrid interface configuration. [16].

The objective of this study was to develop a mathematical model for predicting stroke incidence. To this end, a sample of 367 stroke patients was selected for the purpose of conducting a regression analysis. 500 individuals who had not previously had a stroke and had been monitored for 2 years made up the control group [17]. Based on data pertaining to a number of well-known modifiable risk variables, it created a logistic regression model [18].

In 65% of instances, the proposed model was able to distinguish between healthy participants and stroke patients [19]. By using the model used in this work to predict the likelihood of having a stroke within a year, high-risk individuals' stroke prevention methods may be improved [20]. The method used to create the algorithm for preventing strokes may be utilized to create analogous models for the early identification of other illnesses [21].
Ghosh et al.'s 2023 study focuses on "Water Quality Assessment Through [22] Predictive Machine Learning", highlighting the use of machine learning for analyzing and predicting water quality parameters. In "Unraveling the Heterogeneity of Lower-Grade Gliomas," Rahat, Ghosh, and colleagues (2023) delve [23] into deep learning-assisted segmentation and genomic analysis of brain MR images, offering new insights into this medical condition. Potato Leaf Disease Recognition and Prediction using Convolutional Neural Networks," by Ghosh,[24] Rahat, and team (2023), showcases the application of convolutional neural networks in accurately identifying diseases in potato leaves. Mandava, Vinta, Ghosh, and Rahat's 2023 research presents "An All-Inclusive Machine Learning and Deep [25] Learning Method for Forecasting Cardiovascular Disease in Bangladeshi Population", integrating advanced AI techniques for health predictions. The 2023 study by Mandava [26] et al., titled "Identification and Categorization of Yellow Rust Infection in Wheat through Deep Learning Techniques", applies deep learning methods to detect and categorize wheat infections effectively. Khasim, Rahat, Ghosh [27], and colleagues' 2023 article, "Using Deep Learning and Machine Learning: Real-Time Discernment and Diagnostics of Rice-Leaf Diseases in Bangladesh", explores AI-based solutions for diagnosing rice-leaf diseases. Deciphering Microorganisms through Intelligent Image Recognition", authored by Khasim, Ghosh, Rahat [28] and others in 2023, discusses the use of machine learning and deep learning in identifying microorganisms through advanced image recognition techniques. The 2023 study by Mohanty, Ghosh, Rahat, and Reddy [29], "Advanced Deep Learning Models for Corn Leaf Disease Classification", focuses on the application of deep learning in classifying diseases in corn leaves based on a field study. Alenezi and team's 2021 research, "Block-Greedy and CNN Based Underwater Image Dehazing for Novel Depth Estimation [30] and Optimal Ambient Light", investigates novel CNN-based methods for enhancing underwater image clarity and depth estimation.

The present investigation was employed to formulate a health risk assessment algorithm for the purpose of forecasting the likelihood of stroke. The profile outlines various risk factors that may contribute to the occurrence of stroke. These factors include age, systolic blood pressure, the use of antihypertensive medication, diabetes mellitus, smoking, prior cardiovascular disease.[31].

The Cox proportional hazards model was employed to estimate the likelihood of stroke, utilizing data from 472 stroke incidents that transpired during a 10-year monitoring period, spanning biannual examinations 9 and 14. The risk factors present in an individual's profile, which can be easily detected during a routine physical examination, may be utilized to compute the likelihood of experiencing a stroke [32].

The identification of individuals with significantly elevated stroke risk and the implementation of multifactorial risk factor modification may be facilitated by the detection of borderline levels of multiple risk factors, including moderate or borderline hypertension [33].

## 3. Existing System

Ayurveda remedies are often used by stroke patients, although there are no published data on their effectiveness or safety. Ayurvedic remedies are not very accurate at diagnosing stroke illness. Through the use of machine learning algorithms, this accuracy may be increased [25]. Until they have one, people cannot foresee getting one or their risk of having one. Very few applications use clinical data that already exists to make predictions, and even those that do are limited by the many link rules that are already in place. The doctor's professional opinion and the patient's medical background are the only things that can be used to figure out the issue.

## 4. Proposed System

A forecast system for illness awareness is absolutely necessary to cope with an issue. The area of artificial intelligence known as machine learning, which gets its training from real-world occurrences, offers excellent assistance in anticipating any form of

event. Machine learning methods such as Logistic Regression, Random Forest, and SVM are employed in the suggested system to make predictions. To choose the approach that offers the greatest accuracy for the dataset, train the models and determine their accuracy. This device may determine how probable it is that a person will get sick by looking at their medical background, including their age, blood pressure, sugar levels, and other things. Classification algorithms are used to predict illness when there are a lot of things to consider.

## 5. Working Methodology

The SVM model, which was trained with the dataset, is applied to the data provided to the GUI for stroke risk prediction, and the user's new data is compared to the trained model. As depicted in Figure, this enables us to determine the optimal method for receiving user input during GUI implementation.

### 5.1 Homepage


**Figure 1.** Predictor Homepage

### 5.2 Stroke

This GUI presents the user with an HTML form where, after entering the necessary information, the user may determine if they are at risk of having a brain stroke or not.

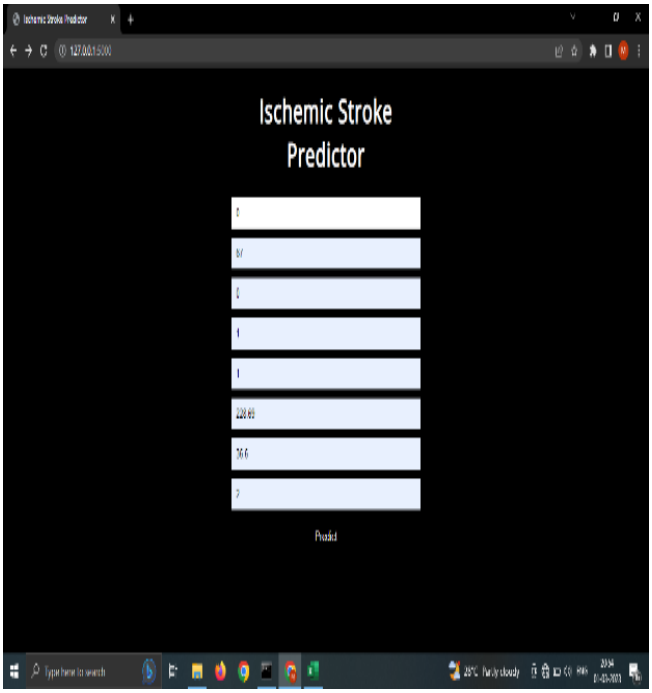


**Figure 2.** Stroke Condition Prediction

The following page will appear if the user is expected to have no stroke, i.e., if the user is determined to not be at risk of having a stroke.

### 5.3 No Stroke


**Figure 3.** No stroke Prediction

This GUI presents the user with an HTML form where, after entering the necessary information, the user may determine if they are at risk of having a brain stroke or not. The following page will appear if the user is projected to have a stroke, or is determined to be at risk of developing one:

This GUI presents the user with an HTML form where, after entering the necessary information, the user may determine if they are at risk of having a brain stroke or not. The following page will appear if the user is projected to have a stroke, or is determined to be at risk of developing one:


**Figure 4.** Predictor Output

## 6. Conclusion

The suggested approach uses a few user-provided inputs to predict correct outcomes with the use of trained Machine Learning algorithms, which helps to forecast brain strokes in a cost-effective and efficient manner. The project is focused on developing a machine learning model that evaluates three algorithms (Decision Tree, SVM, and Random Forest) and uses one of the three to deploy to a flask application in order to detect patients who are at risk for strokes. Therefore, the system for predicting brain strokes has been constructed utilising the three machine learning algorithms listed above, with the maximum accuracy being 94.30%. In order to provide a user interface that is both straightforward and effective while also showing empathy for both users and patients, the system was created in this manner. Utilising a prediction model, it assists in predicting the risk of stroke and offers individualized warnings and advice on changing one's lifestyle. A number of preventative measures, such as quitting smoking and abstaining from alcohol, are recommended to reduce the risk of having a stroke.

## 7. Future Scope

Future expansion of the system has the potential to provide better outcomes and an improved user experience. As a consequence, the user will save significant time and be better prepared to act appropriately in light of the outcomes. The installed system's potential future use may include improving the model's accuracy.

To improve the results, it's important to get a sample with enough traits and more data. More study needs to be done to improve the current methods to machine learning so that they can work in real time and make a good model. Additionally, the performance of the generated model has to be evaluated using data sets with various data volumes.

## References


[1] Kanna, R. K., Devi, K. Y., Dhivy, A. J. A., Amutha, P. M. D., Gomalavalli, R., & Ambikapathy, A. (2022, December). Software Development Framework for Cardiac Disease Prediction Using Machine Learning Applications. In 2022 IEEE International Conference on Current Development in Engineering and Technology (CCET) (pp. 1-7). IEEE.

[2] Li, C., Liu, X., Shen, P., Sun, Y., Zhou, T., Chen, W., ... & Gao, P. (2023). Improving Cardiovascular Risk Prediction through Machine Learning Modelling of Irregular Repeated Electronic Health Records. European Heart Journal-Digital Health, ztad058.

[3] Amagada, P. U. (2023, October). An Inferable Machine Learning Approach for Reservoir Lithology Characterization Using Drilling Data. In SPE Annual Technical Conference and Exhibition? (p. D011S999R007). SPE.

[4] Sun, W., Fang, Y., Wang, J., Yin, D., Chao, X., Zhang, F., ... & Liu, X. (2023). Post-stroke apathy biotypes and their relation to the effort-based reward decision network: a resting-state fMRI study.

[5] Gutiérrez-Serafín, B., Andreu-Perez, J., Pérez-Espinosa, H., Paulmann, S., & Ding, W. (2024). Toward assessment of human voice biomarkers of brain lesions through explainable deep learning. Biomedical Signal Processing and Control, 87, 105457.

[6] Raj, R., Kannath, S. K., Mathew, J., & Sylaja, P. N. (2023). AutoML accurately predicts endovascular mechanical thrombectomy in acute large vessel ischemic stroke. Frontiers in Neurology, 14.

[7] Lee, J., Kwon, K., Soltis, I., Matthews, J., Lee, Y., Kim, H., ... & Yeo, W. H. (2023). Intelligent upper-limb exoskeleton using deep learning to predict human intention for sensory-feedback augmentation. arXiv preprint arXiv:2309.04655.

[8] Kanna, R. K., Kripa, N., & Vasuki, R. Systematic Design Of Lie Detector System Utilizing EEG Signals Acquisition. International Journal of Scientific & Technology Research, 9, 610-2.

[9] Sharma, N., Mangla, M., Iqbal, M. M., & Mohanty, S. N. (2023). Deep Learning Framework for Identification of Skin Lesions. EAI Endorsed Transactions on Pervasive Health and Technology, 9(1).

[10] Mondal, S., Ghosh, S., & Nag, A. (2023). Brain stroke prediction model based on boosting and stacking ensemble approach. International Journal of Information Technology, 1-10.

[11] Mohanty, S. N., Diaz, V. G., & Kumar, G. S. (Eds.). (2023). Intelligent Systems and Machine Learning: First EAI International Conference, ICISML 2022, Hyderabad, India, December 16-17, 2022, Proceedings, Part I (Vol. 470). Springer Nature.

[12] Kanna, R. K., Mutheeswaran, U., Ramya, V. S., Gomalavalli, R., Hema, L. K., & Ambikapathy, A. (2022, December). Computing Model for Alzheimer Prediction Using Support Vector Machine Classifier. In 2022 IEEE International Conference on Current Development in Engineering and Technology (CCET) (pp. 1-6). IEEE.

[13] Kanna, R.K., Ishaque, M., Panigrahi, B.S., Pattnaik, C.R. (2024). Prediction of Covid-19 Using Artificial Intelligence [AI] Applications. In: Roy, B.K., Chaturvedi, A., Tsaban, B., Hasan, S.U. (eds) Cryptology and Network Security with Machine Learning. ICCNSML 2022. Algorithms for Intelligent Systems. Springer, Singapore.

[14] Singh, R., Sharma, H. K., Choudhury, T., Mor, A., Mohanty, S., & Mohanty, S. N. (2023). Blockchain for IoT-enabled Healthcare. EAI Endorsed Transactions on Pervasive Health and Technology, 9.

[15] Rahman, A., Rahman, F. B. A., Islam, A., Jahan, I., & Salam, K. M. A. (2023, July). Cerebral Stroke Prediction Using Machine Learning Algorithms. In 2023 3rd International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME) (pp. 1-7). IEEE.

[16] Kanna, R. K., Gomalavalli, R., Devi, Y., & Ambikapathy, A. (2023). Computational Cognitive Analysis for Intelligent Engineering Using EEG Applications. In Intelligent Engineering Applications and Applied Sciences for Sustainability (pp. 309-350). IGI Global.

[17] Kanna, R. K., Surendhar, S. P. A., AL-Hameed, M. R., Lafta, A. M., Khalid, R., & Hussain, A. (2023, May). Smart Prosthetic Arm Using Cognitive Application. In 2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE) (pp. 1330-1334). IEEE.

[18] Bajaj, K., & Jain, R. (2023, July). Teeth Staining from Tobacco Consumption as an Indicator to Chronic Illness: A Data Analytics and Machine Learning Application. In 2023 World Conference on Communication & Computing (WCONF) (pp. 1-7). IEEE.

[19] R. K. Kanna, V. Subha Ramya, A. A. Khafel, K. A. Jabbar, M. Al-Tahee and R. Khalid, "Cognitive Disability Prediction & Analysis using Machine Learning Application," 2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India, 2023, pp. 1359-1364.

[20] Naveen, S., Upamanyu, M. S., Chakki, K., Chandan, M., & Hariprasad, P. (2023, July). Air Quality Prediction Based on Decision Tree Using Machine Learning. In 2023 International Conference on Smart Systems for applications in Electrical Sciences (ICSSES) (pp. 1-6). IEEE.

[21] Ravikumar, K. K., Ishaque, M., Panigrahi, B. S., & Pattnaik, C. R. (2023). Detection of Covid-19 Using AI Application. EAI Endorsed Transactions on Pervasive Health and Technology, 9.

[22] Ghosh, H., Tusher, M.A., Rahat, I.S., Khasim, S., Mohanty, S.N. (2023). Water Quality Assessment Through Predictive Machine Learning. In: Intelligent Computing and Networking. IC-ICN 2023. Lecture Notes in Networks and Systems, vol 699. Springer, Singapore. https://doi.org/10.1007/978-981-99-3177-4_6

[23] Rahat IS, Ghosh H, Shaik K, Khasim S, Rajaram G. Unraveling the Heterogeneity of Lower-Grade Gliomas: Deep Learning-Assisted Flair Segmentation and Genomic Analysis of Brain MR Images. EAI Endorsed Trans Perv Health Tech [Internet]. 2023 Sep. 29 [cited 2023 Oct. 2];9.https://doi.org/10.4108/eetpht.9.4016

[24] Ghosh H, Rahat IS, Shaik K, Khasim S, Yesubabu M. Potato Leaf Disease Recognition and Prediction using Convolutional Neural Networks. EAI Endorsed Scal Inf Syst [Internet]. 2023 Sep. 21https://doi.org/10.4108/eetsis.3937

[25] Mandava, S. R. Vinta, H. Ghosh, and I. S. Rahat, “An All-Inclusive Machine Learning and Deep Learning Method for Forecasting Cardiovascular Disease in Bangladeshi Population”, EAI Endorsed Trans Perv Health Tech, vol. 9, Oct. 2023.https://doi.org/10.4108/eetpht.9.4052

[26] Mandava, M.; Vinta, S. R.; Ghosh, H.; Rahat, I. S. Identification and Categorization of Yellow Rust Infection in Wheat through Deep Learning Techniques. EAI Endorsed Trans IoT 2023, 10. https://doi.org/10.4108/eetiot.4603

[27] Khasim, I. S. Rahat, H. Ghosh, K. Shaik, and S. K. Panda, "Using Deep Learning and Machine Learning: Real-Time Discernment and Diagnostics of Rice-Leaf Diseases in Bangladesh", EAI Endorsed Trans IoT, vol. 10, Dec. 2023 https://doi.org/10.4108/eetiot.4579

[28] Khasim, H. Ghosh, I. S. Rahat, K. Shaik, and M. Yesubabu, "Deciphering Microorganisms through Intelligent Image Recognition: Machine Learning and Deep Learning Approaches, Challenges, and Advancements", EAI Endorsed Trans IoT, vol. 10, Nov. 2023. https://doi.org/10.4108/eetiot.4484

[29] Mohanty, S.N.; Ghosh, H.; Rahat, I.S.; Reddy, C.V.R. Advanced Deep Learning Models for Corn Leaf Disease Classification: A Field Study in Bangladesh. Eng. Proc. 2023, 59, 69.https://doi.org/10.3390/engproc2023059069

[30] Alenezi, F.; Armghan, A.; Mohanty, S.N.; Jhaveri, R.H.; Tiwari, P. Block-Greedy and CNN Based Underwater Image Dehazing for Novel Depth Estimation and Optimal Ambient Light. Water 2021, 13, 3470. https://doi.org/10.3390/w13233470

[31] Rahman, A., Hoque, S., Khan, M. T., Rahman, F. B. A., Islam, A., Shumona, S. I., & Salam, K. M. A. (2023, June). Early Prediction of Ischemic Stroke Using Machine Learning Boosting Algorithms. In 2023 3rd International Conference on Intelligent Technologies (CONIT) (pp. 1-7). IEEE.

[32] Kanna, R. K., & Vasuki, R. (2019). Advanced Study of ICA in EEG and Signal Acquisition using Mydaq and Lab view Application. International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN, 2278-3075.

[33] Kaur, R., & Dhillon, S. (2023). Brain Stroke Prediction System–A Review. International Journal of Modern Developments in Engineering and Science, 2(6), 21-24.

[34] Kanna, R. K., Ramya, V. S., Khafel, A. A., Jabbar, K. A., Al-Tahee, M., & Khalid, R. (2023, May). Cognitive Disability Prediction & Analysis using Machine Learning Application. In 2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE) (pp. 1359-1364). IEEE.