



AUTOMATED STROKE PREDICTION USING MACHINE LEARNING

Zakeer N¹, Dr C Rangaswamy², Harianirudh, Hariprasad³

Dept of Electronics and Communication, SJC Institute of Technology, Chickballapur, India

ABSTRACT:

Stroke is a leading cause of disability and mortality worldwide, emphasizing the critical need for accurate prediction models to identify individuals at high risk. Machine learning techniques have shown promise in this domain due to their ability to analyze complex data sets and extract meaningful patterns. This paper presents a comprehensive approach to automated stroke prediction using machine learning algorithms. The proposed model incorporates diverse features including demographic information, medical history, lifestyle factors, and clinical indicators to capture a holistic view of stroke risk. Feature selection techniques are employed to identify the most informative predictors, enhancing the model's efficiency and interpretability.

Keywords: Machine learning, Stroke prediction, Analyze data, Risk factors, Data pre processing, Model training, Model evaluation.

Introduction:

Burden of stroke in the World- Stroke is the second leading cause of death and leading cause of adult disability worldwide with 400-800 strokes per 100,000, 15 million new acute strokes every year, 28,500,000 disability adjusted life-years and 28-30 day case fatality ranging from 17% to 35%. The burden of stroke will likely worsen with stroke and heart diseases related deaths projected to increase to five million in 2020, compared to three million in 1998. This will be a result of continuing health and demographic transition resulting in increase in vascular disease risk factors and population of the elderly. Developing countries account for 85% of the global deaths from stroke. The social and economic consequences of stroke are substantial. The cost of stroke for the year 2002 was estimated to be as high as \$49.4 billion in the United States of America (USA), while costs after discharge were estimated to 2.9 billion Euros in France. Causes of mortality from stroke-Death from stroke is as a result of comorbidities and / or complications. Complications of stroke may arise at different time periods.

LITERATURE SURVEY:

Paper 1

Title: "Computer Methods and Programs in Biomedicine"

Authors : Jae-woo Lee, Hyun-sun Lim, Dong-wook Kim, Soon-aeShin, Jinlwon Kim , BoraYoo, Kyung-hee Cho.

Published on : 17 Jan 2021

Description : The Purpose of this paper was calculation of 10-year stroke prediction probability of stroke into five categories.

Paper 2

Title : "Probability of Stroke : A Risk Profile from the Framingham Study"

Authors: Philip A. Wolf, MD; Ralph B.D'Agostino, PhD, Albert J. Belanger, MA; and William B. Kannel.

Published on: 16 Feb 2024

Description : In this paper, A health risk appraisal function has been developed for the prediction of stroke using the Framingham Study cohort.

Paper 3

Title : "Development of an Algorithm for Stroke Prediction: A National Health Insurance Database Study"

Authors : Min SN , Park SJ, Kim DJ, Subramaniam M , Lee KS

Published on : 20 Nov 2023

Description : In this research , this paper aimed to derive a model equation for developing a stroke pre diagnosis algorithm with the potentially modifiable risk factors.

Paper 4

Title : "Stroke prediction using artificial intelligence".

Authors : M.Sheetal Singh, Prakash Choudhary

Published on : 15 May 2023

Description : In this paper, Here, decision tree algorithm is used for feature selection process, principal component analysis algorithm is used for reducing the dimension and adopted back.

Machine learning in stroke prediction:

Machine learning algorithms are trained on data on patients and their medical histories, as well as information about their risk factors and results, in the context of stroke prediction . The objective is to create models that can properly forecast a patient's chance of having a stroke, and then utilize that knowledge to identify individuals who are at high risk and take preventive measures. ML algorithms can examine vast volumes of data and uncover patterns and correlations that would be impossible to notice by hand. These models' results can be utilized to enhance diagnosis, therapy, and patient outcomes.

How Does Machine Learning Technology Work ?

Stroke prediction using machine learning involves the utilization of advanced computational algorithms to analyze various risk factors and predict the likelihood of an individual experiencing a stroke within a certain time frame. Here's an overview of how it works :

1. **Collection:** The process begins with gathering relevant data from diverse sources, including electronic health records (EHRs), medical imaging, patient demographics, lifestyle factors, genetic information. These datasets provide valuable insights into the patient's health status and potential risk factors associated with stroke.
2. **Feature Selection:** Machine learning models require input features that are indicative of stroke risk. Feature selection involves identifying and extracting relevant variables from the collected data, such as age, blood pressure, cholesterol levels, smoking status, medical history and pre-existing conditions like diabetes or atrial fibrillation.
3. **Data Preprocessing:** Raw data often contain noise, missing values, and inconsistencies that can affect the performance of machine learning algorithms. Data preprocessing techniques such as normalization, imputation of missing values, and feature scaling are applied to ensure the data is clean and standardized for analysis.
4. **Model Training:** Once the data is prepared, it is divided into training and testing sets. The training set is used to train machine learning on historical data, allowing it to learn patterns and relationships between input features and stroke outcomes. Various machine learning algorithms, such as logistic regression, decision trees, random forests, support vector machines, or deep learning neural networks, can be employed for this task.
5. **Model Evaluation:** After training, the model's performance is evaluated using the testing set to assess its accuracy, sensitivity, specificity, and other relevant metrics. Cross-validation techniques may be used to ensure the robustness of the model and prevent overfitting, where the model performs well on training data but fails to generalize to new data.
6. **Prediction:** Once the model is trained and validated, it can be deployed to predict stroke risk for new patients. Input data containing the relevant features of a patient are fed into the trained model, which then generates a probability or classification indicating the likelihood of the individual experiencing a stroke within a specified timeframe (e.g., next 5 years).

ADVANTAGES

1. Guide treatment decisions.
2. Data Accuracy.
3. Efficiency.
4. Highly accurate guesses based on historical data.
5. linear regression technique.

APPLICATIONS

1. Predict likelihood of stroke
2. Emergency Alert.
3. Health Monitoring.
4. Safety during Outdoor Activities.

Conclusion

Stroke prediction using machine learning represents a significant advancement in the field of healthcare, offering a promising avenue for more accurate and personalized risk assessment. Through the integration of complex algorithms with diverse patient data, machine learning models have demonstrated remarkable capabilities in forecasting the likelihood of stroke occurrence.

By leveraging sophisticated techniques for data analysis and pattern recognition, these models can identify subtle correlations and risk factors that may not be apparent through traditional methods. This enhanced predictive accuracy enables healthcare professionals to stratify patients based on their individual risk profiles, facilitating targeted interventions and preventive strategies.

However, challenges such as data quality, model interpretability and ethical considerations must be addressed to ensure the responsible and equitable implementation of machine learning in stroke prediction. By harnessing the predictive prowess of machine learning, we can empower healthcare providers, optimize patient care, and ultimately work towards a future where strokes are not just treated, but prevented altogether.

REFERENCES:

1. "Computer Methods and Programs In the Biomedicine"- Jae-woo Lee, Hyun-sun Lim, Dong-wook Kim, Soon-ae Shin, Jinkwon Kim, Bora Yoo, Kyung-hee Cho
2. "Probability of Stroke: ARisk Profile from the Framingham Study' -Philip A.Wolf, MD; Ralph B.D'Agostino, PhD,Alberr J.Belanger,MA; and William B.Kennal,MD
3. Development of an Algorithm for Stroke Prediction: A National Health Insurance Database Study" -Min SN, Park SJ, Kim DJ, Subramaniam M, Lee KS
4. Stroke prediction using artificial intelligence".-M.Sheetal Singh, Prakash Choudhary.
5. T.Elloker and A.J.Rhoda, "The relationship between social support and participation in stroke: A systematic review," Afr.J.Disability, vol.7,pp. 1-9, Oct.2018.