IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

PREDICTION OF BRAIN STROKE USING MACHINE LEARNING TECHNIQUE

Miss. K. BHAVANI [1], Mrs. R.Vanithamani [2], Dr. D. RAJINIGIRINATH [3]

PG Scholar [1], Department of Computer Science, SRI Muthukumaran Institute of Technology, India
Assistant Professor [2], Department of Computer Science, SRI Muthukumaran Institute of Technology, India
HOD (CSE/AIDS) [3], Department of Computer Science, SRI Muthukumaran Institute of Technology, India

ABSTRACT

Cerebral stroke is a life-threatening medical condition that requires immediate diagnosis and intervention to minimize its devastating effects. Timely and accurate classification of stroke subtypes is crucial for determining appropriate treatment strategies and improving patient outcomes. In this research, we explore the application of supervised learning approaches to classify cerebral stroke subtypes based on relevant medical data. The study involves the collection of a diverse dataset comprising anonymized patient records, including clinical indicators, medical history, risk factors, and diagnostic imaging results. The data is carefully pre-processed to handle missing values, normalize numerical features, and address any potential biases. The experimental results demonstrate the effectiveness of supervised learning approaches in accurately classifying cerebral stroke subtypes. The model with the highest performance metrics is identified, and its potential integration into clinical settings is discussed. Additionally, the research explores potential challenges and limitations of the proposed approach, along with potential strategies to improve model accuracy and generalizability.

EXISTING SYSTEM

Neural-like P systems are membrane computing models inspired by natural computing and are viewed as third-generation neural network models. Although real neurons have complex structures, classical neural-like P systems simplify the structures and corresponding mechanisms to two-dimensional graphs or tree-based firing and forgetting communications, which limit the real applications of these models. In this paper, we propose a hypergraph-based numerical neural-like (HNN) P system containing five types of neurons to describe the high-order correlations among neuron structures. Three new kinds of communication mechanisms among neurons are also proposed to address numerical variables and functions. Based on the new neural-like P system, a tumor/organ segmentation model for medical images is developed. The experimental results indicate that the proposed models outperform the state-of-the-art methods based on two hippocampal datasets and a multiple brain metastases dataset, thus verifying the effectiveness of the HNN P system in correctly segmenting tumors/organs.

DISADVANTAGES

- **They did not implement the deployment process.**
- **They only disease segmentation.**
- **Accuracy & performance was low.**
- **They did not using machine learning techniques.**

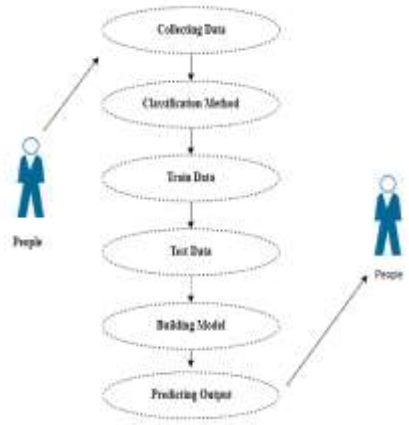
PROPOSED SYSTEM

The proposed system aims to enhance network security by predicting and pre-emptively detecting various types of network attacks using supervised machine learning techniques. The system leverages a diverse dataset of network traffic, comprising both normal and attack instances, to develop an efficient predictive model. The first step involves collecting a comprehensive dataset of network traffic. This dataset should include normal network traffic as well as instances of various attack types, such as DDoS attacks, intrusion attempts, malware activity, etc. The data is then pre-processed to handle missing values, normalize numerical features, and encode categorical variables. Next, a set of relevant features is extracted from the pre-processed data to represent the characteristics of network traffic. Feature selection techniques, such as Recursive Feature Elimination (RFE) or Principal Component Analysis (PCA), are applied to reduce dimensionality and select the most informative features, thereby improving the efficiency and effectiveness of the predictive model. Several supervised machine learning algorithms are considered for the predictive model. Popular algorithms like Decision Trees, Random Forests, Support Vector Machines (SVM), Neural Networks, and K-Nearest Neighbours (KNN) are evaluated and compared based on their performance metrics using cross-validation techniques.

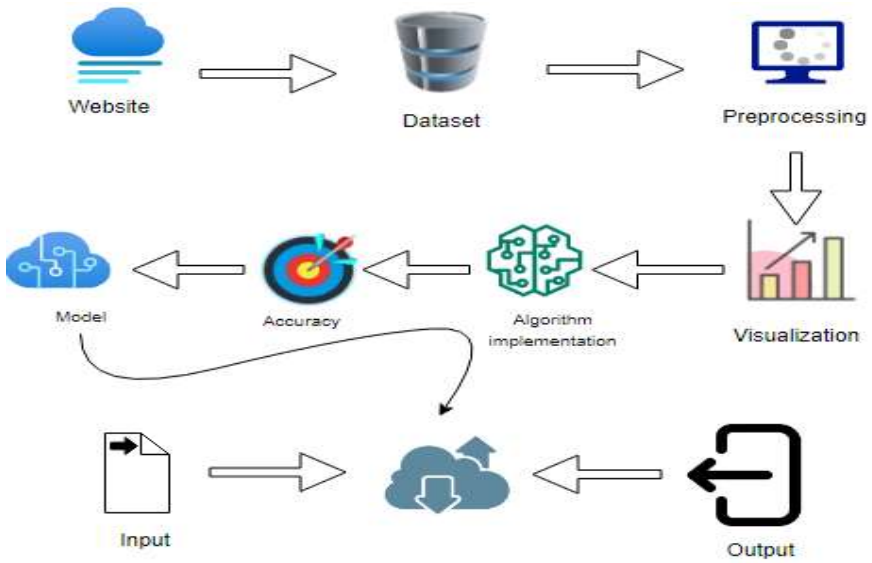
ADVANTAGES

- We build framework based application for deployment purposes.
- We using machine learning techniques to build a predictive model.
- Accuracy & performance level improved.
- We implemented the Data analysis process by using histograms, Plots, and Graphs.
- We compared more than two algorithms to get a better accuracy level.

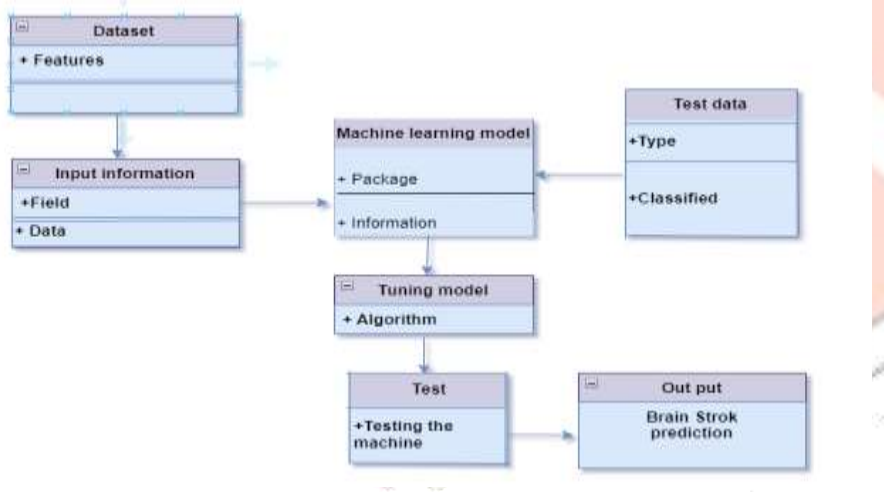
USE CASE DIAGRAM



SYSTEM ARCHITECTURE



CLASS DIAGRAM



DATA PRE PROCESSING

Validation techniques in machine learning are used to get the error rate of the Machine Learning (ML) model, which can be considered as close to the true error rate of the dataset. If the data volume is large enough to be representative of the population, you may not need the validation techniques. However, in real-world scenarios, to work with samples of data that may not be a true representative of the population of given dataset. To finding the missing value, duplicate value and description of data type whether it is float variable or integer.

DATA VISUALIZATION

Data visualization is an important skill in applied statistics and machine learning. Statistics does indeed focus on quantitative descriptions and estimations of data. Data visualization provides an important suite of tools for gaining a qualitative understanding. This can be helpful when exploring and getting to know a dataset and can help with identifying patterns, corrupt data, outliers, and much more. With a little domain knowledge, data visualizations can be used to express and demonstrate key relationships in plots and charts that are more visceral and stakeholders than measures of association or significance. Data visualization and exploratory data analysis

are whole fields themselves and it will recommend a deeper dive into some the books mentioned at the end.

FUTURE WORK

- The added background knowledge from other datasets can also possibly improve the accuracy of stroke prediction models as well.
- We intend to collect our institutional dataset for further benchmarking of these machine learning methods for stroke prediction.
- We also plan to perform external validation of our proposed method, as a part of our upcoming planned work.

CONCLUSION

After the literature survey, we came to know various pros and cons of different research papers and thus, proposed a system that helps to predict brain strokes in a cost effective and efficient way by taking few inputs from the user side and predicting accurate results with the help of trained Machine Learning algorithms. Thus, the Brain Stroke Prediction system has been implemented using the given Machine Learning algorithm given a Best accuracy.The system is therefore designed providing simple yet efficient User Interface design with an empathetic approach towards their users and patients.

REFERENCES

Naghavi, M.; Abajobir, A.A.; Abbafati, C.; Abbas, K.M.; Abd-Allah, F.; Abera, S.F. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: A systematic analysis for the Global Burden of Disease Study 2016. Lancet 2017, 390, 1151–1210. [Google Scholar] [CrossRef] [Green Version]
2. Hay, S.I.; Abajobir, A.A.; Abate, K.H.; Abbafati, C.; Abbas, K.M.; Abd-Allah, F. Global, regional, and national disability-adjusted life-years (DALYs) for 333 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. Lancet 2017, 390, 1260–1344. [Google Scholar] [CrossRef] [Green Version]
3. Vos, T.; Abajobir, A.A.; Abate, K.H.; Abbafati, C.; Abbas, K.M.; Abd-Allah, F. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. Lancet 2017, 390, 1211–1259. [Google Scholar] [CrossRef] [Green Version]
4. Wang, W.; Jiang, B.; Sun, H.; Ru, X.; Sun, D.; Wang, L. Prevalence, Incidence, and Mortality of Stroke in China. Circulation 2017, 135, 759–771. [Google Scholar] [CrossRef] [PubMed]
5. Zhou, M.; Wang, H.; Zeng, X.; Yin, P.; Zhu, J.; Chen, W. Mortality, morbidity, and risk factors in China and its provinces, 1990–2017: A systematic analysis for the Global Burden of Disease Study 2017. Lancet 2019, 394, 1145–1158. [Google Scholar] [CrossRef] [Green Version]
6. Xing, X.; Yang, X.; Liu, F.; Li, J.; Chen, J.; Liu, X. Predicting 10-Year and Lifetime Stroke Risk in Chinese Population. Stroke 2019, 50, 2371–2378. [Google Scholar] [CrossRef]
7. Feigin, V.L.; Nguyen, G.; Cercy, K.; Johnson, C.O.; Alam, T.; Parmar, P.G. Global, Regional, and Country-Specific Lifetime Risks of Stroke, 1990 and 2016. N. Engl. J. Med. 2018, 379, 2429–2437. [Google Scholar] [CrossRef]
8. Wang, Y.; Liu, J.; Wang, W.; Wang, M.; Qi, Y.; Xie, W. Lifetime risk of stroke in young-aged and middle-aged Chinese population: The Chinese Multi-Provincial Cohort Study. J. Hypertens. 2016, 34, 2434–2440. [Google Scholar] [CrossRef] [Green Version]
9. Abedi, V.; Goyal, N.; Tsivgoulis, G.; Hosseinichimeh, N.; Hontecillas, R.; Bassaganya-Riera, J. Novel Screening Tool for Stroke Using Artificial Neural Network. Stroke 2017, 48, 1678–1681. [Google Scholar] [CrossRef]
10. Doupe, P.; Faghmous, J.; Basu, S. Machine Learning for Health Services