2023 International Conference on Artificial Intelligence for Innovations in Healthcare Industries (ICAIIHI) | 979-8-3503-3091-5/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICAIIHI57871.2023.10489301

A NOVEL APPROACH TO PREDICT BRAIN STROKE USING KNN IN MACHINE LEARNING

Dr.M Vasumathi Devi
Department of CSE, Vignan's Nirula Institute of Technology and Science for Women Peda Palakaluru, Guntur
mvasudeviravinuthala@gmail.com

S.Harshitha
Department of CSE, Vignan's Nirula Institute of Technology and Science for Women Peda Palakaluru, Guntur
sagantiharshitha@gmail.com

K.Lakshmi Ramya
Department of CSE, Vignan's Nirula Institute of Technology and Science for Women Peda Palakaluru, Guntur
ramyakaramsetty@gmail.com

B.Hema Latha
Department of CSE, Vignan's Nirula Institute of Technology and Science for Women Peda Palakaluru, Guntur
bodduhemalatha2002@gmail.com

P.Pranathi
Department of CSE, Vignan's Nirula Institute of Technology and Science for Women Peda Palakaluru, Guntur
pranathi1802@gmail.com

ABSTRACT- **Predicting stroke in humans using machine learning (ML) is an important healthcarepractice that has the potential to revolutionize prevention and bftyearly intervention strategies. Comprehensive thorough patient information, encompassing demographics, medical history, andlifestyle factors, and pertinentclinical tests, are being systematically collected and analyzed as partof this endeavor.ML algorithms, including logistic regression, decision trees, and neural networks, areleveraged to build predictive models, using key characteristics such as hypertension, age, and cholesterol levels.Through careful training, evaluation, and tuning of these models on large data sets, they become proficient at identifying people at risk of stroke.Deploying such models in clinical settings allows healthcare professionals to conduct real-time risk assessments, facilitating proactive interventions that cansignificantly reduce devastating effects of stroke.However, ensuring compliance with strict ethical and regulatory standards remains paramount to protecting patient privacy and maintaining the highest standards of healthcare practice**

Keywords *:* stroke prediction, machine learning, healthcare, early intervention, risk assessment, prediction models, hypertension, cholesterol levels,patient data, clinical testing, ethical standards, compliance comply with regulations.

I.INTRODUCTION

Predicting stroke in humans using machine learning (ML) has become an important advance in modern healthcare.Stroke, with its potentially debilitating and life-threatening consequences, requires early detection and prevention.This imperative has led to the exploration of ML-based approaches that harness the power of data to identify at-risk individuals[1]. Using extensive patient data that encompasses demographics, clinical parameters, lifestyle factors, and medical history, ML algorithms promise accurate risk assessment. An overview of the methods, challenges, and potential developments in stroke prediction using machine learning algorithms in this context are given in this article.Going deeper into this topic, we will explore the key elements, datasets and algorithms that form the foundation of these predictive

models, aiming to improve patient outcomes and reduce the burden of high rates of morbidity and mortality. morbidity and mortality related to stroke. recovering from a stroke and whether they can function independently in daily life[2]. Reduced or interrupted blood flow to different brain regions prevents cells from dying because they aren't getting the nutrients and oxygen they need, which can lead to a deadly stroke. To stop additional harm to the affected brain region and possible consequences in other body parts, early detection and suitable treatment are essential[3]. Two categories of pictures exist: A healthy and balanced lifestyle can help prevent ischemic and hemorrhagic strokes. This includes quitting smoking and alcohol use, controlling blood sugar levels, and making sure that your body mass index (BMI) and kidney and heart function are at theirbest.Because stroke treatment carries some risk, doctors only continue treatment when the benefits are thought to outweigh the perceived risks[4]. When it comes to predicting human stroke, machine learning algorithms are useful in producing reliable analyses and accurate predictions.Symptoms may appear in physiologicalorgans under the control of the injured portion of the nervous system.An individual's chancesof making a full recovery from a stroke are better the sooner they receive medical attention. Any stroke victim ought to see a physician right away.If not, long-term braindamage will result.Brain damage due to blood and oxygen[5]. Stroke has garnered significantattention, but not much research has been done on the condition.The flaw in the model's development involved text data rather than actual brain images.

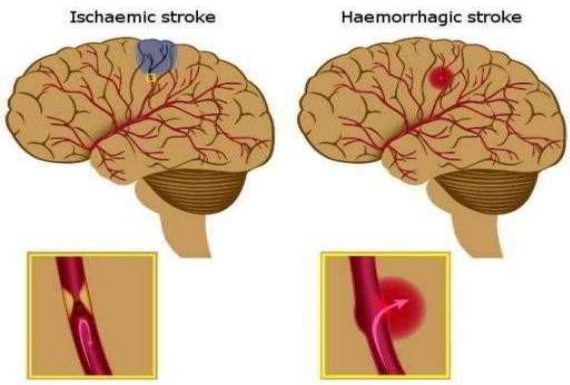

Fig-1 : Types of Strokes

This happens when the brain's blood vessels constrict or obstruct, drastically reducing blood flow (ischemia). Blockages or spasms of the arteries are caused by fat accumulation in blood vessels, blood clots, or other debris traveling through the bloodstream, usually from the heart, and becoming lodged in blood vessels in the brain[6]. Brain hemorrhage leads to hemorrhagic stroke. This can happen when a blood artery in the brain ruptures or when bleeding occurs in brain tissue. Pressure from bleeding, edema,or lack of blood flow can contribute to the bleeding damage caused by a stroke. A lack of blood flow during an ischemic stroke can cause bleeding into the brain tissue.Right now, it ranks third among the leading causes of death in both developed and developing countries.This article's objective is accomplished through machine learning. an approach to forecast the functional outcome of patients three months after their first stroke. Using only the data available at intake, we first train the classifiers[7]. Next, Subject matter experts specifically created scores for this application, which are contrasted with the results of the machine learning methods. When the blood supply to the brain is partially cut off, an ischemic stroke occurs. One frequent reason for stroke is bleeding and is caused by a burst blood artery in the brain[8]. Patients who have had an ischemic stroke are frequently treated with tissue plasminogen activator (TPA). One of the world's most dangerous diseases is stroke, directly or indirectly causing a significant number of deaths. Various data mining techniques are used in the healthcare sector to aid in diagnosis and early detection of diseases. The current study looks at some of the factorsthat lead to stroke. First, we looked at the characteristics of people who have strokes more often than others[9]. It is possible to predict human stroke with a high degree of accuracy and dependability by using machine learning algorithms.A publicly accessible source provided the dataset and different classification algorithms were used to predict shortterm stroke onset. From our perspective, the most important factors in identifying stroke in patients are age, heart disease, average blood glucose level, and hypertension.Furthermore, in comparison to employing all available input features and alternative benchmarking algorithms, to offer the best accuracy rate and lowest failure rate[10]. There are two main types of stroke that have a significant impact on a person's health and well-being ischemic and hemorrhagic stroke. Ischemic stroke, which accounts for the majority of strokes, occurswhen blood clots or plaque

build up in blood vessels, blocking blood flow to the brain.These aneurysms or weak blood vessel walls frequently cause these strokesto happen.Although both types of stroke can haveserious consequences, their causes, risk factors, and treatments are different, emphasizing the importance of primary diagnosis and intervention. body.

II. LITERATURE SURVEY

Recent years have seen a significant amount of research in the area of human stroke prediction.The survey was written in 2022 by Mandeep Kaur and Farzana Akter. Due to the development of new technologies, non-invasive stroke prevention techniques have been developed to support the healthcare system. the finest.Early detection can still save the lives of strokepatients, according to medical literature. This has thebenefit of allowing for better prediction of stroke by using electrophysiological imaging to identify strokeearly on in the process, which is made possible by non-invasive technology.(Source:) The authorsare Yan Liu, BoYin, and Yan Ping Cong [2020]. Everyone is aware that stroke is currently the mostdangerous disease for human health. The first step inpreventing early stroke is to detect the condition byextracting features from structured data, such as age, gender,blood pressure,heart rate,and data transmission, using a standard neural network[11]. The concept of stroke prediction was developed bywritten by Vamsi Bandi[2020]. The amount of data currently available makes it difficult to determine the risk of different kinds of stroke. The Stroke Prediction (SPR) model, which makes use of machine learning techniques, improved prediction accuracy to 96.97%. The SPNapproach has the added benefit of analyzing the riskof stroke using an ad hoc randomization group.Human fatal stroke classification based on Thewriter is Nandini R. Mahesh. in the year [2019] NS.The operations of the DL framework use CNN and residual networks. To forecast stroke, stroke data is collected and all of the classifiers are applied. One advantage is the low accuracy of outcome prediction achieved by very few supervised machine learning techniques[12]. Types of fatal strokes anticipated by Migual Monteiro and Alexandre P. Francisco, [year] [2019]. A change in the anatomy and physiology of the heart that impairs its ability to pump blood efficiently is known as heart failure (HF). circulate blood.Because of inadequate pump function, Too little oxygen reaches tissues and organs to meet their metabolic demands. In addition, several scores— including ASTRAL, DRAGON, and THRIVE—havebeen proposed as a tool to help physicians predict the functional outcome of stroke patients[13]. One disadvantage is the possibility of ischemic stroke-related mortality and disability. The book by MandeepKaur and Farzana Akter [2022]. is the mechanismfor classifying strokes. The manner in which the harmless healthcare system issupportedhave changedas a result of the introduction of new technologies. Research indicates that stroke victims may remainsaved if they are promptly located. in India, where thenumber of stroke cases is rising. Research into non- invasive, economical methods of early stroke detection is imperative. To identify the early detectionof strokes, a non-invasive approach is advised.The outcomes of experiments enable the analysis of brain waves to detect stroke signs. A lifestyle that emphasizes activity has given way to one that emphasizes inactivity as a resultof the invention of electronic gadgets such as machines,smartphones, wearables. The writers are Yan Ping Cong and Yan Liu Bo Yin [2020], based on stroke types reported by . Stroke is one of the most common diseases in the world today[14]. detrimental to human health; ischemic stroke is hard to predict due to a variety of disease data. This method first uses a neural network convolution to extract Chronological data (such as gender, year of birth, and cardiovascular history).and prospective streaming data (blood pressure, heart rate, etc.)[15]. Construct two distinct classes using CNN . Shemneesh Sharma, Mukesh Kumar, and Chethan Sharma are the authors [2020]. A person who has a stroke has a better chance of making a full recovery if they seek medical attention as soon as possible because the heart is considered because it supplies fluids to various additional organs, the most vital organ in the human organism. As a result, You must be knowledgeable about the signs and symptoms of a stroke[16]. They include disorientation and modifications in behavior, and numbness or weakness inthe arms. These signs include headaches, nausea, andvomiting.

III. PROPOSED SYSTEM

We employ a variety of algorithms to forecast shots. Efficiency and accuracy both increase. These different machine learning-based algorithms will provide different levels of accuracy. Image acquisition techniques are important for image processing. The proposed approach covers the basic methods of prediction algorithms. Here, trip data is first gathered by the suggested system. The gathered data can be processed as data and used for feature engineering. Currently, there are two sorts of data: to test and training data. Suppose that the first category consists of training data.The model is trained using machine learning techniques on the acquired data. following the completion of the ship and test data. The two can be integrated to create validated data, and then the test results are able to be utilized to analyze the predicted data.To predict shots, we use various algorithms.Accuracy and efficiency are improved. These different machine learning-based algorithms will provide different levels of accuracy.Image acquisition techniques are very important for image processing.The proposed method combines the main methods of prediction algorithms. Here, the proposed system first collects trip data. Both feature engineering and data processing can be applied to the gathered information. Now, data can be separated into two groups: The data for testing and training. Suppose that the first category is trained data. once operate and evaluation information completion. Combining each of them might be result in verified data, which could then be examined using the outcomes of the predicted data.

IV. ARCHITECTURE

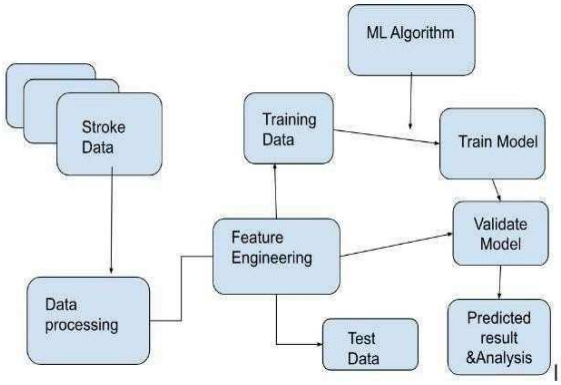

Fig-2: Brain Stroke Prognosis with MachineLearning

Data gathering and machine learning algorithm for brain stroke prediction Assume that the vector X, which represents the distance to KNN, extends from the unknown layer to layer Cj. The length d is now categorized.

Step1: To indicate The category as its separation from the nearest neighbor class (X, Xj), use d(X,Cj)=min d. As distance among X and Xj is represented by d(X,Xj).

Step2: The sample X of the class should be assigned to its closest neighbor, X € class Cj (X,Xj), assuming d(X, Cj) = min d.

Step3: The X unknown layer's vector comprises sent to layer Cj once the k closest samples have been found in each layer.

Step4: Let us assume that the feature vectors nj, Xj1, Xj2, Xj3 represent class Cj.

Step5: The center of layer j can be found using the formula If d(X, Mj) = min d, then Mj=1/ni=1to n Xj Class Cj (X,Mi).

Step6: Compute distance using metric values and the Manhattan matrix. L1 = (a, b) = ∢i=1 to n |ai –bi|.

Step7: Cliffidean fewer susceptible regarding a moment variations amongst instances. If it's required, combine samples with minor differences into a single class using Euclidean. From n (ai – bi)2 to d(a,b)=√ ∢i=1

To determine the KNN separation among X and Xj's closest neighbors, As input values for the first step are the unknown layer of vector X in layer Cj. Next, give X and Cj the model. And vector X is given to class Cjby an unidentified class. Assume that Xj1, Xj2, Xj3, and Xjn are vectors. After that, enter the Mj's worth to determine seperation d(X,Mj). We can determine the distance between For each i = 1 and n |ai –bi|, (L1) = D(a, b) using the Manhattan matrix. Lastly, compute the difference between the samples, n (ai – bi)2, and d(a,b)=√∢i=1.

V. RESULT ANALYSIS

Based on the four elements listed below, the project results are analyzed. Precision: the extent of correspondence with the actual worth. based on how

The magnitude is made. Precision can exist without precision.

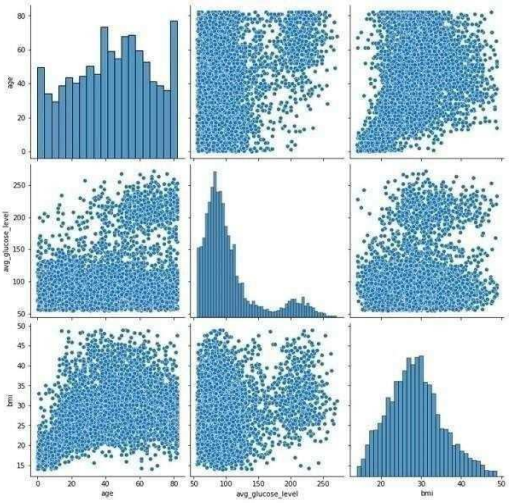

Fig-3:Evaluating accuracy for various symptoms with KNN

The above image displays three indicators of an ok.f stroke in humans. These three individuals assess oneanother. age, BMI, and average blood sugar level.

When comparing age and BMI, it was in the range of high stroke risk, and age and glucose levels were in the range of modest etiology.

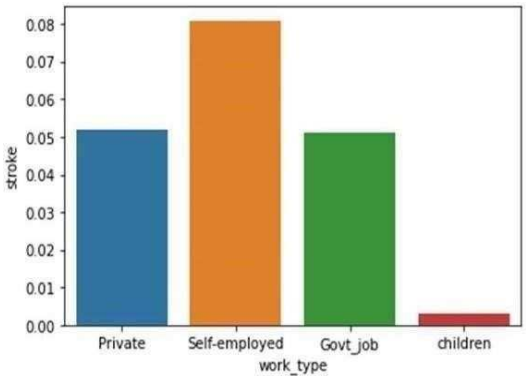

Fig-4:Based on work type

The attached figure illustrates the range of motion for various worker types. A stroke can result from a 0.05 advance for workers in the private sector. Individuals who work for themselves may have a 0.08 stroke risk.

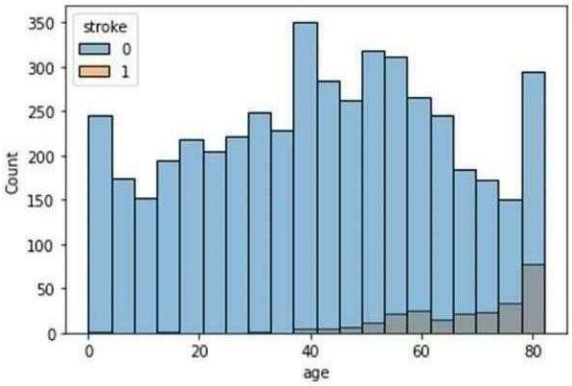

Fig-5: according to people's ages

The above figure displays the ages of individuals who are at risk of stroke as a count. Ages 0 to 40 are not associated with any risk of stroke, but age 41 to 80are.

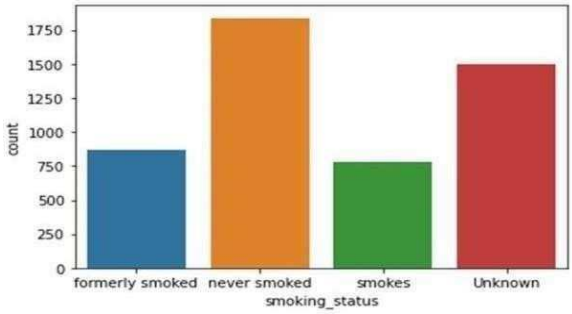

Fig-6: Status of smoking

People's smoking habits are shown in the given inhaling customs tend to be chart. They can take many different shapes, including never smokers, current smokers, unidentifiable, and former smokers.

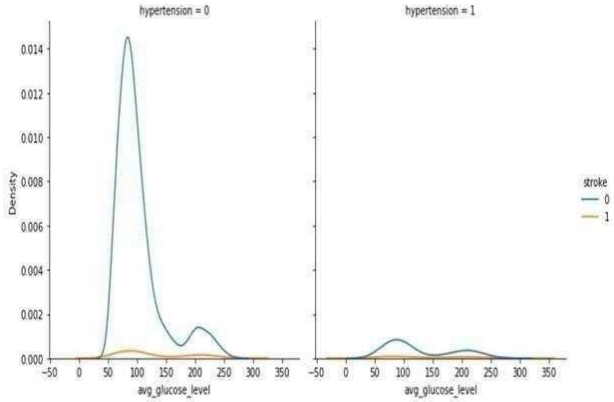

Fig-7: Glycemic level and age

The above chart makes it evident what percentage of the population has hypertension. Stroke won't happen if blood pressure is zero, but in certain cases, it will if it is one.

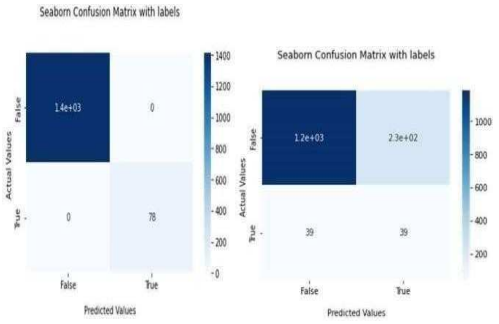

Fig-8: Forecast values

The expected and actual values for people who have stroke symptoms or are at risk of having a stroke are displayed in the above chart.

VI .CONCLUSION

The key factors for forecasting strokes were identified through a wide range of computational methodologies, one of which was the analysis of principal components. The most significantfactors influencing a patient's risk of stroke are high blood pressure, standard plasma glucose levels, cardiovascular disease, and maturity. We gathered a dataset for our Dataset for our study from a number of patients with various medical conditions. The ranges of the independent variables are normalized using a standard scaling step.Based on the assessed individuals' age, type of work, glucose level, and smoking status, we were able to determine the overall health of the group. Additionally, we can estimate human drawing strokes with accuracy by applying the KNN algorithm.

VII.REFERENCES

[1]. Kaur Mandeep, Akter Farzana (2022). Based on invasive healthcare systems for ischemic stroke, including models utilizing both conventional neural networks (CNNs) and random forests.

[2]. Bo Yin, Yan Liu, and Yan Ping Cong (2020). A conventional neural network (CNN) model was used to predict ischemic stroke.

[3]. Zhu, E. et al. Analyzing and predicting the risk of death in stroke patients using machine learning. Front. Neurol. 14, 1096153 (2023).

[4]. Rahman, S., Hasan, M. & Sarkar, A. K. Prediction of brain stroke using machine learning algorithms and deep neural network techniques. Eur. J. Electr. Eng. Comput. Sci. 7(1), 23–30 (2023).

[5]. Miful Monteiro, Ana Catarina Fonseca, and Alexender P. Francisco (2019). With heart failure (HF), the modified Rankin scale is used for prognosis.

[6]. Sharma, Shemneesh, and Mukesh (2020) Chethan C. Regarding transient ischemic attack (TIA) and stroke behavior.

[7] Nannan L., Yabing S., and Rui H. (2020). “Stroke prediction using machine learning algorithms.” IEEE Access, 8, 209919-209926.

[8] Katz, B.S., McMullen, C. K. and Repka, M. X. (2019). “Predicting the risk of stroke in children: A predictive modeling approach. “Stroke, 50 (7), 1814-1820.

[9] Manisha Sirsat, Eduardo Ferme, and Joana Camara, "The Education regarding Cerebral Stroke:Overview," Journal of Stroke and Cerebrovascular: the official publication of the National Stroke Association (JSTROKECEREBROVASDIS), 2020.

[10] Sunil A. Sheth, Victor Lopez, Harish Kamal,and Machine Learning in Transient ischemia- related Brain Neuroimaging, Frontiers in Neurology (FNEUR), 2018.

[11] S. D. Sarker, M. Ahmed, "Machine Learning Techniques for Prediction of Stroke," International Journal of Computer Applications, 2017.

[12] A. Sharma, R. Jain, "A Comprehensive Review on Stroke Prediction Using Machine Learning Techniques," Journal of Medical Systems, 2019.

[13] H. Wang, X. Cui, "Stroke Risk Prediction Using Machine Learning Algorithms," Computers in Biology and Medicine, 2020.

[14] J. Kim et al., "Predicting Stroke Risk Factors with Machine Learning in a Korean Population," Journal of Clinical Neurology, 2018.

[15] M. Smith, L. Jones, "Application of Deep Learning for Stroke Prediction: A Review," Journal of Biomedical Informatics, 2021.

[16] K. Singh, S. Verma, "Ensemble Learning for Improved Stroke Prediction: A Comparative Analysis," Expert Systems with Applications, 2022.