Literature Review

* introduction

According to the World Health Organization (who), stroke is the second leading cause of death in the world, accounting for about 11% of all deaths. Stroke is a kind of sudden cerebrovascular event. If we can predict ahead of time, we can take effective preventive measures for high-risk groups, which is of great significance in reducing the incidence rate of stroke. Prediction model is the main means to predict the risk of stroke. At present, most of the existing stroke early warning tools at home and abroad are based on the traditional mathematical model of risk factors and simple early warning tools. Because of its complex calculation, its clinical application is limited. In this study, according to the information data set of stroke patients provided by medical institutions, through the analysis of the characteristics of the data, we build a stroke prediction model, accurately predict the onset time of patients, and help medical staff to take timely protective measures.

* Literature Review

Stroke, also known as "brain attack", is one of the major causes of disability and death around the world. According to the World Health Organization (who), stroke is the second leading cause of death in the world, accounting for about 11% of all deaths. In the United States, there are nearly 800000 new stroke patients every year - about 137000 people die as a result, and the survivors permanently change their lifestyles. However, if the occurrence of early stroke can be determined or predicted as soon as possible, 80% of stroke can be prevented [1] [2]. It is estimated that there are currently 6.5 million stroke survivors in the United States. Stroke is a kind of brain damage caused by the interruption or reduction of blood supply to the brain. The oxygen and nutrition supply to the brain is cut off, and brain cells begin to die within a few minutes. Therefore, stroke is considered as an emergency medical event, which needs immediate diagnosis and treatment. Accurate judgment of the root cause and location of stroke is the decisive factor for the therapeutic effect. Many years ago, people thought that stroke could not be cured, but now, we have been able to try new technologies to treat stroke. The progress of medical technology has greatly improved the accuracy of stroke diagnosis and brain injury diagnosis. Prediction model is the main means to predict the risk of stroke [3] [4] [5] [6] [7]. According to the research of Akash, kunder [8] and others, using deep learning method to build stroke prediction model has achieved good results.

According to the Research Report of kurmi Sima [9], it is not enough for patients to know the early warning symptoms, risk factors and acute stroke reaction of stroke. Therefore, it is scientific and necessary to rely on the third party to predict the stroke of patients. At present, most of the existing stroke early warning tools at home and abroad are based on the traditional mathematical model of risk factors and simple early warning tools [10]. Because of its complex calculation, its clinical application is limited. In this study, according to the information data set of stroke patients provided by medical institutions, by analyzing the characteristics of the data, we built a stroke prediction model, accurately predicted the onset time of patients, and helped medical staff to take timely protective measures [11] [12].

According to the research of Nakano K [13], the risk factors of stroke include hypertension, smoking, high cholesterol and overweight, age, gender, race, family history, cardiovascular disease, recurrent stroke, TIA, etc. [14]. Stroke indicators: (1) high blood pressure: high blood pressure and blood pressure continuously higher than 115 / 75, high blood pressure can increase the risk of stroke by 2 to 6 times; (2) smoking: smoking for more than 5 years can double the risk of stroke; (3) high cholesterol: high cholesterol can increase the risk of stroke, and the optimal level of cholesterol in adults is 200 or 200 (4) overweight: overweight increases the risk of stroke, and weight should be related to height, so it is not easy to be overweight; (5) age: Although young people may have stroke, after 55 years old, the risk of stroke will double every ten years [15]; (6) gender: the risk of stroke in men is slightly higher than that in women. (7) Race: African Americans almost double the risk of stroke compared to Caucasians. (8) Family history: people with a family history of stroke have a higher risk of stroke. (9) Cardiovascular disease: heart attack can increase the risk of stroke, atrial fibrillation (AF) arrhythmia can also increase the risk of stroke [16] [17]; (10) recurrent stroke: people with a history of stroke have a higher risk of recurrent stroke. This high risk will last about five years and then gradually decrease; the risk is highest in the first few months. (11) TIA: the risk of severe stroke increases in the month of transient ischemic attack (TIA), usually within two days. And it is extremely easy to cause death after stroke because IHM related factors include diabetes [18] [19] [20].

According to N. Stroebele [21] studies have shown that stroke has early warning signals, and includes sudden and intermittent symptoms, such as acute numbness, weakness or paralysis on the face, arms or legs, especially on one side of the body; inability to raise both arms above the head at the same time, or inability to smile normally; sudden loss of balance, dizziness, loss of coordination; sudden blurred vision and decreased vision in one or both eyes Or see things with double shadow; sudden confusion of consciousness, difficulty in speaking or understanding simple sentences; sudden unexplained, severe local headache, sometimes accompanied by vomiting[22] [23]. The more symptoms you have, the more likely you are to have a stroke. Every minute after a stroke warning signal appears is very important, ideally within 60 minutes. Researchers have found that brain damage associated with stroke is more harmful than the damage directly caused by stroke and can worsen within 24 hours. According to the data of Zeng Huangling and other researchers, the recurrence of stroke will leave sequelae, such as spasticity [24], as the medical staff said, "loss of time is loss of brain".

If we can use tools to accurately identify, correctly judge these symptoms, record the time of these symptoms and other information, and then build a prediction model through deep learning, we can accurately judge the current physical condition of patients, predict whether patients will have a stroke in a certain period of time, and send them to the hospital for treatment in time, so as to reduce the possibility of permanent injury .

1. Summary

This paper refers to 30 literatures at home and abroad, and has a certain understanding of the theoretical knowledge and application methods of stroke prediction. This paper summarizes the characteristics and problems of the current stroke prediction methods.

Aiming at the problem of low accuracy of stroke prediction methods, this paper adopts the idea of ensemble learning, uses a series of learners to learn, and uses some rules to integrate the learning results, so as to obtain significantly superior generalization performance than a single learner. Because the problems that a single machine learning model can solve are limited and the generalization ability is poor, but by building and combining multiple learners to complete the learning task, we can often get the miraculous effect. These learners can be regarded as a basic unit, and they can be combined to form a powerful whole, which can solve more complex problems. Its idea can be summarized as three parts The cobbler is better than Zhuge Liang [25] [26] [27] [28] [29].

Advantages of integrated learning:

1. There are some differences between individual learners, which will lead to different classification boundaries, that is, there may be errors. After merging multiple individual learners, we can get more reasonable boundaries, reduce the overall error rate and achieve better results.

2. When the data set is too large or too small, it can be divided and put back to generate different data subsets, and then use the data subsets to train different learners, and finally merge into a strong learner;

3. If the boundary of data partition is too complex and it is difficult to describe the situation with linear model, then multiple models can be trained and then fused.

4. For multiple heterogeneous feature sets, it is difficult to fuse them directly, so we can consider using each data set to build a classification model, and then fuse multiple models.

A total of 5111 data sets were used in this experiment. The data sets included gender, age, hypertension, heart disease, marriage, work style, residence type, average blood glucose level, body mass index, smoking status, etc. Considering the small data set and some missing values, this paper intends to use three representative methods of ensemble learning framework, which are random forest method using bagging framework, AdaBoost method using boosting framework and stacking model method using our data to train the three models, and compare the prediction accuracy and recall of the three models The most suitable model for stroke prediction was selected. By using different combination strategies and finding the optimal parameters, the classification performance of the model is further improved.

* References

[1] “Probability of Stroke: A Risk Profile from the Framingham Study” -Philip A. Wolf, MD; Ralph B. D'Agostino, PhD,Albert J. Belanger, MA;and William B.Kannel,MD

[2] “Focus on stroke: Predicting and preventing stroke” - Michael Regnier

[3] Knowledge and application of upper limb prediction models and attitude toward prognosis among physiotherapists and occupational therapists in the clinical stroke setting[J]. Topics in Stroke Rehabilitation, 2021, 28(2) : 135-141.

[4] Chiu IMin et al. Using a Multiclass Machine Learning Model to Predict the Outcome of Acute Ischemic Stroke Requiring Reperfusion Therapy[J]. Diagnostics, 2021, 11(1) : 80-80.

[5] Vascular Diseases and Conditions - Ischemia; Investigators at Ocean University of China Zero in on Ischemia (The Probability of Ischaemic Stroke Prediction with a Multi-Neural-Network Model)[J]. Information Technology Newsweekly, 2020, : 318-.

[6] Development and validation of a 2-year new-onset stroke risk prediction model for people over age 45 in China.[J]. Medicine, 2020, 99(41) : e22680-e22680.

[7] Kamilla Steensig et al. A Novel Model for Prediction of Thromboembolic and Cardiovascular Events in Patients Without Atrial Fibrillation[J]. The American Journal of Cardiology, 2020, 131 : 40-48.

[8] Akash, Kunder & Shashank, H & .S, Srikanth & A.M, Thejas. (2020). Prediction of Stroke Using Machine Learning.

[9] Khedr Eman M. et al. Post-stroke dysphagia: frequency, risk factors, and topographic representation: hospital-based study[J]. The Egyptian Journal of Neurology, Psychiatry and Neurosurgery, 2021, 57(1)

[10] Chen YunHsuan, Sawan Mohamad. Trends and Challenges of Wearable Multimodal Technologies for Stroke Risk Prediction. 2021, 21(2):460-460.

[11] Dardick Joseph M. et al. Associating cryptogenic ischemic stroke in the young with cardiovascular risk factor phenotypes[J]. Scientific Reports, 2021, 11(1)

[12] “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

[13] Nakano K K. An overview of stroke. Epidemiology, classification, risk factors, clinical aspects.[J]. Postgraduate medicine, 1986, 80(4) : 82-8, 93, 96-7.

[14] Klissurski M. et al. A rare case of non-conventional risk factors for ischemic stroke: A long pathogenic sequence, neurointervention combined with IV acetylsalicylic acid and secondary prevention[J]. Interdisciplinary Neurosurgery: Advanced Techniques and Case Management, 2021, 23.

[15] Krzystanek Ewa et al. Adequate Knowledge of Stroke Symptoms, Risk Factors, and Necessary Actions in the General Population of Southern Poland.[J]. Brain sciences, 2020, 10(12)

[16] Benz Alexander P, Healey Jeff S, Chin Ashley, et al. Stroke risk prediction in patients with atrial fibrillation with and without rheumatic heart disease.2021.

[17] Kurmi Sima et al. Awareness of Stroke Warning Symptoms, Risk Factors, and Response to Acute Stroke in Biswanath District, Assam, India[J]. Journal of Stroke Medicine, 2020, 3(2) : 88-91.

[18] Vamsi Bandi, Debnath Bhattacharyya, Divya Midhunchakkravarthy. Prediction of Brain Stroke Severity Using Machine Learning. 2020, 34(6)

[19] Huang ZhiXin et al. Risk factors for in-hospital mortality among acute ischemic stroke patients in China: a nationwide prospective study.[J]. Neurological research, 2020, : 1-9.

[20] Establishment of Clinical Prediction Model Based on the Study of Risk Factors of Stroke in Patients With Type 2 Diabetes Mellitus.[J]. Frontiers in endocrinology, 2020, 11 : 559-559.

[21] N. Stroebele et al. Systematic review of women's knowledge of warning signals and risk factors of stroke[J]. European Journal of Integrative Medicine, 2009, 1(4) : 230-230.

[22] Hausfater Pierre. [Nosocomial heat stroke: warning signal or forgotten victims of August 2003 heat wave?].[J]. Presse medicale (Paris, France : 1983), 2006, 35(2 Pt 1) : 194-5.

[23] Busse O. [Transient ischemic attacks as a warning signal. As stroke precursor it is often underestimated].[J]. Fortschritte der Medizin, 1995, 113(30) : 431-3.

[24] Zeng Huangling et al. Prevalence and Risk Factors for Spasticity After Stroke: A Systematic Review and Meta-Analysis [J]. Frontiers in Neurology, 2021,

[25] “Development of an Algorithm for Stroke Prediction: A National Health Insurance Database Study” - Min SN, Park SJ, Kim DJ, Subramaniyam M, Lee KS

[26] “Stroke prediction using artificial intelligence”- M. Sheetal Singh,Prakash Choudhary

[27] “Medical software user interfaces,stroke MD application design (IEEE)” -Elena Zamsa

[28] “Effective Analysis and Predictive Model of Stroke Disease using

Classification Methods” - A.Sudha, P.Gayathri,N.Jaisankar

[29] “Deep learning algorithms for detection of critical findings in head CT scans: a retrospective study” - Rohit Ghosh,Swetha Tanamala, Mustafa Biviji,Norbert G Campeau, Vasantha Kumar Venugopal

[30] M. M. Mirbagheri,W. Z. Rymer.Predication of Reflex Recovery After Stroke Using Quantitative Assessments of Motor Impairment at 1 Month[C].//2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2009). [v.20].2009:7252-7255.

[31] Gorelick, P.B.; Qureshi, S.; Farooq, M.U. Management of blood pressure in stroke. Int. J. Cardiol. Hypertens. 2019, 3, 100021.

[32] Rastegar, S.; GholamHosseini, H.; Lowe, A. Non-invasive continuous blood pressure monitoring systems: Current and proposed

technology issues and challenges. Phys. Eng. Sci. Med. 2020, 43, 11–28. 48. Arakawa, T. Recent Research and Developing Trends ofWearable Sensors for Detecting Blood Pressure. Sensors 2018, 18, 2772.

[33] El-Hajj, C.; Kyriacou, P.A. A review of machine learning techniques in photoplethysmography for the non-invasive cuff-less

measurement of blood pressure. Biomed. Signal Process. Control 2020, 58, 101870.

[34] Hosanee, M.; Chan, G.;Welykholowa, K.; Cooper, R.; Kyriacou, P.A.; Zheng, D.; Allen, J.; Abbott, D.; Menon, C.; Lovell, N.H.; et al.

Cuffless Single-Site Photoplethysmography for Blood Pressure Monitoring. J. Clin. Med. 2020, 9, 723.

[35] Kario, K.; Saito, I.; Kushiro, T.; Teramukai, S.; Tomono, Y.; Okuda, Y.; Shimada, K. Morning Home Blood Pressure Is a Strong Predictor of Coronary Artery Disease: The HONEST Study. J. Am. Coll. Cardiol. 2016, 67, 1519–1527.

[36] Lin,W.-H.; Zhang, H.; Zhang, Y.-T. Investigation on Cardiovascular Risk Prediction Using Physiological Parameters. Comput. Math.Methods Med. 2013, 2013, 272691.

[37] Asayama, K.; Ohkubo, T.; Kikuya, M.; Obara, T.; Metoki, H.; Inoue, R.; Hara, A.; Hirose, T.; Hoshi, H.; Hashimoto, J.; et al.Prediction of Stroke by Home “Morning” Versus “Evening” Blood Pressure Values. Hypertension 2006, 48, 737–743.

[38] Liu, F.-D.; Shen, X.-L.; Zhao, R.; Tao, X.-X.; Wang, S.; Zhou, J.-J.; Zheng, B.; Zhang, Q.-T.; Yao, Q.; Zhao, Y.; et al. Pulse pressure as an independent predictor of stroke: A systematic review and a meta-analysis. Clin. Res. Cardiol. 2016, 105, 677–686.

[39] Ecobici, M.; Stoicescu, C. Arterial Stiffness and Hypertension—Which Comes First? Maedica (Buchar) 2017, 12, 184–190.

[40] Castaneda, D.; Esparza, A.; Ghamari, M.; Soltanpur, C.; Nazeran, H. A review on wearable photoplethysmography sensors and

their potential future applications in health care. Int. J. Biosens. Bioelectron. 2018, 4, 195–202.

[41] Tang, Y.;Wang, M.Y.;Wu, T.T.; Zhang, J.Y.; Yang, R.; Zhang, B.; Shi, Y.; Meng, P.; Ji, N.; Sun, Y.; et al. The role of carotid stenosis ultrasound scale in the prediction of ischemic stroke. Neurol. Sci. 2020, 41, 1193–1199.