MEDS Public Health and Preventive Medicine (2023)
Clausius Scientific Press, Canada

DOI: 10.23977/phpm.2023.030605
ISSN 2616-1915 Vol. 3 Num. 6

Prognostic study of hemorrhagic stroke patients based on GABP neural network model

Anqi Wang[1,2,3,*], Jing Zhao[1], Dongchen Mao[1,2,3], Bo Li[1,2,3], Mingke Li[2,3]

[1]School of Electronic Information, Xijing University, Xi'an, China
[2]Henan Ruika Robot Manufacturing Co., LTD, Xinxiang, China
[3]Shanghai Shuoka Robot Technology Co., LTD, Shanghai, China
[*]Corresponding author: 18134402844@163.com

Keywords: GABP, mRS, Prediction Model, Hemorrhagic stroke

Abstract: Hemorrhagic stroke is a serious cerebrovascular disease, which seriously endangers our life safety. Clinical intelligent diagnosis and treatment of hemorrhagic stroke, as a combination of artificial intelligence and intelligent medical treatment, is very conducive to improving the monitoring of patients' pathological cycle changes through data flow. Therefore, based on the BP neural network model, this paper is based on the personal medical history, medical history, morbidity and treatment history of 100 patients with hemorrhagic stroke, as well as patient prognosis assessment information. A neural network model based on GABP was constructed to successfully predict the 90-day mRS Score of patients, and provide recommendations for clinical decision-making according to the correlation between patient prognosis and personal history, medical history, treatment and imaging features.

1. Introduction

Hemorrhagic stroke is a bleeding disorder in the brain that is usually not caused by an external injury, but by causes such as high blood pressure and hardening of the arteries in the brain. Hemorrhagic stroke, also known as cerebral congestion, is a common disease with a high rate of disability and mortality. In recent years, advances in the field of medical imaging and the wide application of artificial intelligence technology have provided new solutions for processing large amounts of medical image data. To construct an intelligent diagnosis and treatment model to predict the risk factors of poor prognosis in hemorrhagic stroke patients and the possibility of cerebral congestion in advance. This will help improve survival and quality of life for patients with hemorrhagic stroke.

Hemorrhagic stroke is a bleeding disorder in the brain that is usually not caused by an external injury, but by causes such as high blood pressure and hardening of the arteries in the brain. Hemorrhagic stroke, also known as cerebral congestion, is a common disease with a high rate of disability and mortality. In recent years, advances in the field of medical imaging and the wide application of artificial intelligence technology have provided new solutions for processing large amounts of medical image data. This will help improve survival and quality of life for patients with hemorrhagic stroke.

2. Prediction model based on GABP neural network

First, a relationship model was established with the patient's personal history, disease history, morbidity and treatment-related features and first image results as input features, and the 90-day mRS Score as output features. Then, a GABP neural network is constructed and the parameters of BP neural network are optimized by genetic algorithm (GA) to improve the prediction accuracy. Finally, mRS Is predicted based on the tuned GABP neural network. Such a model would help to more accurately predict a patient's 90-day mRS Score.

In addition, based on the constructed GABP neural network, this paper also considers the follow-up data to improve the prediction accuracy of mRS. This will allow the model to make more comprehensive use of follow-up data to make predictions and provide more reliable predictions of 90-day mRS Scores.

In this paper, all factors are tested for Shapiro-Wilk distribution to see if they conform to normal distribution. Then Pearson correlation analysis or Spearman [1] correlation analysis is selected to explore the correlation between these factors according to the distribution of data. Finally, a thermal map was drawn to visualize the correlation, showing the relationship between these factors and prognosis.

2.1. Materials and methods

First, a relationship model was established with the patient's personal history, disease history, morbidity and treatment-related features and first image results as input features, and the 90-day mRS Score as output features. Then, a GABP neural network is constructed and the parameters of BP neural network are optimized by genetic algorithm (GA) to improve the prediction accuracy. Finally, mRS Is predicted based on the tuned GABP neural network. Such a model would help to more accurately predict a patient's 90-day mRS Score.

In addition, based on the constructed GABP neural network, this paper also considers the follow-up data to improve the prediction accuracy of mRS. This will allow the model to make more comprehensive use of follow-up data to make predictions and provide more reliable predictions of 90-day mRS Scores.

In this paper, all factors are tested for Shapiro-Wilk distribution to see if they conform to normal distribution. Then Pearson correlation analysis or Spearman [2] correlation analysis is selected to explore the correlation between these factors according to the distribution of data. Finally, a thermal map was drawn to visualize the correlation, showing the relationship between these factors and prognosis.

2.2. Background of the study area

Predicting 90-day mRS Scores in patients with hemorrhagic stroke has multiple input features, including personal history, disease history, treatment, and imaging features, and requires modeling of complex relationships between these features. Although traditional BP neural network [3] can be used to solve this problem, its training process may be affected by local optimal solutions, resulting in degraded model performance.

In order to cope with the limitations of BP neural network in dealing with complex problems, this study introduced genetic algorithm [4] to optimize the neural network model. Genetic algorithm is a kind of natural heuristic algorithm, which can be used for global optimization problems and help to jump out of the local optimal solution, thus improving the performance of the model. In order to enhance the optimization efficiency of the algorithm, the forward propagation mechanism of BP neural network is also fully used to evaluate the fitness of each individual genetic algorithm, so as to guide the search path of the genetic algorithm more effectively.

3. The establishment of model

By combining genetic algorithm with BP neural network [5], a method named GABP[6] (Global Balanced backpropagation) was proposed in this study to deal with this complex prediction problem, so as to further improve model performance and optimization efficiency. This approach has potential when dealing with complex problems and can be used effectively in multiple application areas such as the medical field. The design framework of the algorithm is shown in the following table.

IGABP represents the weight and threshold of the neural network as a continuous vector in the coding part of the genetic algorithm, which is used to compose the gene expression of an individual. In the part of calculating individual fitness, different from traditional GABP, IGABP starts from principle and uses the process of forward propagation of neural network to directly calculate individual fitness. This means that there is no need for training like traditional BP neural networks, and there is no need to calculate the process of error backpropagation, thus eliminating the amount of computation required for training.

The advantage of this method is to improve the optimization efficiency of the algorithm, reduce the computational cost, and more directly use the structure of the neural network to evaluate the fitness of individuals. As a result, IGABP is able to converge to a suitable solution more quickly, especially for neural network problems that require large-scale training.

3.1. The mRS Score was predicted by considering follow-up results

GABP is a hybrid algorithm that combines genetic algorithms and backpropagation neural networks. Its main goal is to use genetic algorithms to fine-tune the weights and thresholds of neural networks in order to significantly improve the performance of neural networks in solving specific problems. The GABP flow chart designed in this paper is shown in Figure 1.

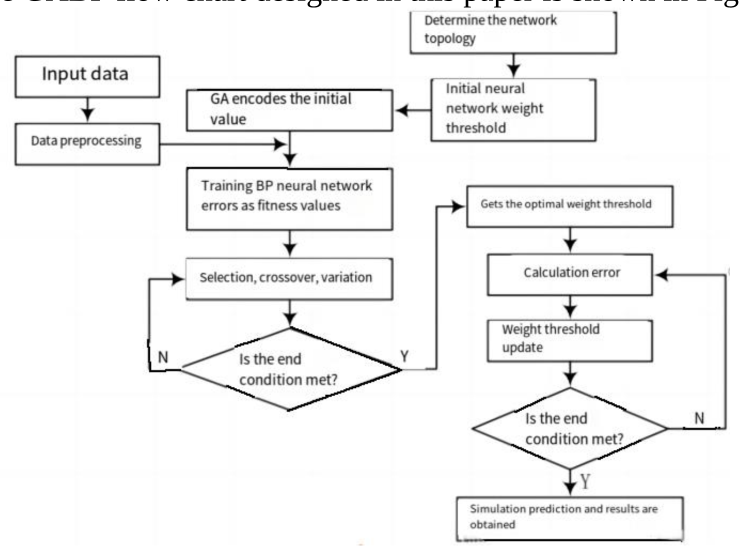

Figure 1: GABP flow chart.

3.2. Model solution results

The predicted values of the first 100 patients were compared with those in Table 1 to obtain the error chart in Figure 2.

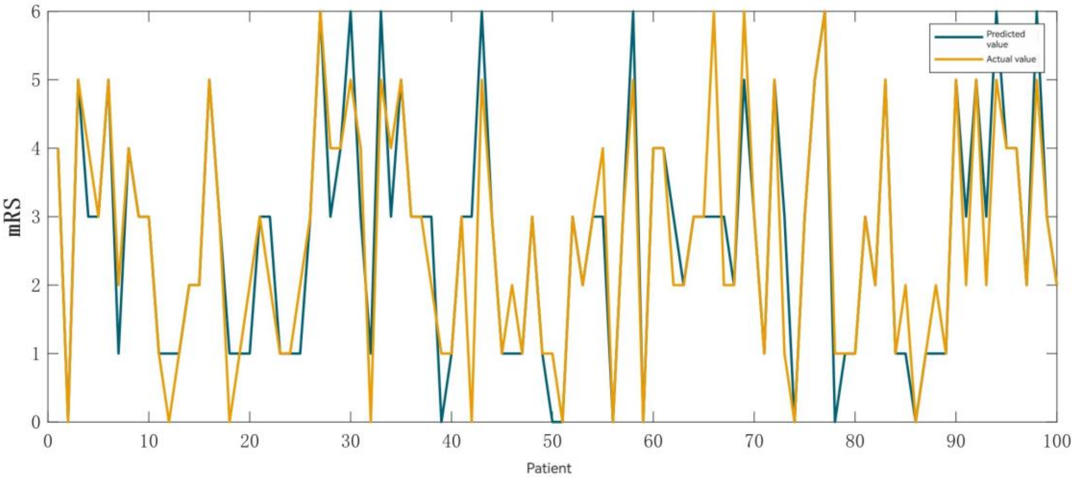

Figure 2: Comparison of predicted values in 100 patients.

3.3. The correlation between prognosis and other factors was analyzed

3.3.1. Establishment of distribution test model

Kolmogorov-Smirnov (KS) distribution test is a statistical test used to test whether data obeys a specific probability distribution (usually a normal distribution) [7]. The null hypothesis (H0) of this test is that the data fits a particular probability distribution, and the alternative hypothesis (H1) is that the data does not.

Let the distribution function of the population X be $F(x)$, $F(x)$ is a continuous function of x,$X_1, X_2, \ldots, X_n$ are the samples from X. Then its cumulative distribution function can be written as:

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I_{[-\infty,x]}(X_i)$$

Where $I_{[-\inf,x]}$ is the indicator function, its value is:

$$\mathrm{I}_{[-\inf,x]}(\mathrm{X_i}) = \begin{cases} 1, \mathrm{X_i} \le \mathrm{x} \\ 0, \mathrm{X_i} > \mathrm{x} \end{cases}$$

At this point, the Kolmogorov-Smirnov statistic can be expressed as:

$$D_n = \sup_x |F_n(x) - F(x)|$$

Where $F_n(x)$ is the cumulative distribution function and $F(x)$ is a hypothetical theoretical distribution. In this paper, we assume that it follows a normal distribution. sup is an upper bound on distance, based on Glivenko-Cantell quantification, if $\mathrm{X_i}$ follows the $F(x)$ theoretical distribution, then D_n approaches 0 as n approaches infinity.

In general, significance is chosen 0.05. If the test value P is less than 0.05, the null hypothesis H_0 is rejected. This indicator does not follow a normal distribution.

The Shapiro-Wilk distribution test model can be described as follows:

The zero test of this test is that the $X_1, X_2, \ldots, X_n$ sample comes from a normally distributed population. The statistics for this test are:

$$W = \frac{\left(\sum_{i=1}^{n} a_i x_{(i)}\right)^2}{\sum_{i-1}^{n} (x_i - \bar{x})^2}$$

$x_{(i)}$ contains the subscript index i in parentheses, not to be confused with x_i, It's a statistic of order i, it's the i smallest number in the sample. $\bar{x}$ is the average of the sample.

Constant a_i is calculated by the following formula:

$$(a_1, \ldots, a_n) = \frac{m^\top V^{-1}}{(m^\top V^{-1} V^{-1} m)^{1/2}}$$

3.3.2. Correlation analysis between disordered variables and 90-day mRS

There are disordered variables in the features to be analyzed, so this part of variables needs to be discussed separately.In general, Kolmogorov-Smirnov chooses 0.05 for significance. If the test P value is less than 0.05 , then the null hypothesis H_0 is rejected, this indicator does not follow a normal distribution. The following takes the relationship with gender as an example.

The results of the normality test [8] are as follows:

Table 1: Normality test list

Index	Significance
The 90-day mRS	2.5293E-19

If the values of P are all less than 0.05, then the null hypothesis H_0 is rejected, and the 90-day mRS Does not obey the normal distribution.The results of variance homogeneity test are shown in Table 2.

Table 2: Test table for homogeneity of variance

	Levin statistics	Significance
The 90-day mRS	289.119480	5.2904E-280

By calculating Levin statistics based on the mean, we find that the significance level is less than 0.1, which indicates that the variance of the data is not uniform. To sum up, for the data of disordered variables, neither the normal distribution nor the hypothesis of homogeneity of variance is satisfied. Therefore, it is appropriate to consider using the Kruskal-Wallis H test to assess the correlation between them.

The analysis results of Kruskal-Wallis H test are shown in Table 3:

Table 3: Kruskal-Wallis H test list

		Daily sales
gender	H	3893.258556
	Prob	0.655502

The results showed that the significance of disorder index was less than 0.05. Therefore, there is no significant correlation between gender and 90-day mRS, that is, gender is not an important factor affecting the 90-day prognosis score of stroke patients. There are several possible explanations and important perspectives for this finding:

(1) Individual differences: Although gender has been recognized as an important predictor in many medical studies, gender does not appear to significantly affect 90-day mRS Scores in this specific population of stroke patients. This suggests that the prognostic score may be influenced by other more important factors, such as age, previous medical history, and severity of disease.

(2) Disease complexity: Stroke is a complex disease, and its prognosis is affected by many factors. While gender can have an impact on the prognosis of stroke patients in some cases, this does not necessarily apply in all cases. Therefore, a comprehensive assessment of prognosis requires a comprehensive consideration of multiple factors.

(3) Statistical methods: Appropriate statistical methods, including Kruskal-Wallis H test, are used to deal with data that do not follow normal distribution and have uneven variance. Such methods allow for more accurate detection of relationships between variables, independent of data distribution.

4. Conclusions

In summary, the GABP neural network model constructed in this paper [9] overcomes the limitations of BP neural network in complex problems, and uses genetic algorithm to optimize the weights and thresholds, thus improving the performance and optimization efficiency of the model. Overall, the GABP-based neural network model represents an approach that integrates neural networks and genetic algorithms and has the potential to solve complex problems. Future research and development will help to further improve the performance, interpretability and applicability of the model, thus promoting its wide application in different fields.

References

[1] Wang Xiaoyan, Li Meizhou, Discussion on grade correlation coefficient and Spearman grade correlation coefficient, Journal of Guangdong Polytechnic of Light Industry, 2006, 4:26-27, 2006

[2] Qi Dehu, Kang Jichang. Design of BP neural network [J]. Computer Engineering and Design, 19 (2): 48-50,1998.

[3] Xuan Guangnan, Cheng Runwei, Genetic Algorithm and Engineering Optimization, China, Tsinghua University Press: 1-10, 2004

[4] Liu Xiufeng, Liu Fen. Classification of colorectal cancer syndrome by BP neural network based on genetic algorithm optimization [J]. Chinese Journal of Medical Library and Information, 27(3): 14-18, 2018. (in Chinese)

[5] Chongruksut W, Limpastan K, Jetjumnong C, et al. Age as a prognostic factor of 30-day mortality in hemorrhagic stroke patients: A Thai large tertiary care referral center[J]. Asian Journal of Surgery, 2020. DOI:10.1016/j.asjsur.2019.11.010.

[6] LIU Chunyan, LING Jianchun, KOU Linyuan. Performance comparison between GA-BP Neural Network and BP Neural Network [J]. Chinese Journal of Health Statistics,(2): 173-176, 2013.

[7] Fan Rong, Meng Dazhi, Xu Dashun, Research Progress of Statistical correlation analysis methods, Mathematical Modeling and Its Application, Vol. 3, No. 1:1-12, 2014

[8] Liang Xiaojun, Normality Test (3), Shanghai Statistics, 2000, 12:24-28

[9] Hu Qiong; Zhai,Xiaoyu; Li,Zhenfu, Multi-Objective Optimization of Deep-Sea Mining Pump Based on CFD, GABP Neural Network and NSGA-III Algorithm, Journal of Marine Science and Engineering, 2022,10:1063-1063