ISSN 1330-3651(Print), ISSN 1848-6339 (Online)
https://doi.org/10.17559/TV-20220814112137
Original scientific paper

A Novel Model for Language Training Assessment Based on Data Mining and Bayesian Network

Jie CHEN

Abstract: At present, the imperfect network language environment has adversely affected cultivation of the language quality. Therefore, it is urgent to use modern information technology to design a novel algorithm model for the cultivation of language quality. To propose a more efficient model for language training assessment, a data mining algorithm and a Bayesian network were used to design a BOPPPS (Bridge-in, Objective/Outcome, Pre-assessment, Participatory learning, Post-assessment, Summary) model, with the model used to cultivate the language quality of adolescents. Regarding adolescents as the research object, a simulation experiment on the BOPPPS model was conducted, evaluating the actual effect by using the multi-level fuzzy comprehensive evaluation model. Results show that the BOPPPS model based on data mining algorithm and Bayesian network can effectively improve the language quality of adolescents. This lifting effect is mainly reflected in the mining and analysis of fMRI brain wave data based on Bayesian network. The resulting BOPPPS model and the traditional simple proposed BOPPPS are more scientific and targeted. During the training of adolescents' language quality, the proposed BOPPPS model can be used directly for teaching applications, which can get a better application effect. The conclusions can provide some reference for the cultivation of language quality and have certain theoretical research significance.

Keywords: Bayesian network; BOPPPS model; cultivation; data mining; fMRI; language quality

1 INTRODUCTION

Since the invention of computers in the middle of the last century, computer and information technology have maintained a rapid development trend, and have brought revolutionary changes and far-reaching effects on human social life [1]. With the development of information technology, the information elements of the three major elements of human society are becoming more and more important, and they are influencing the direction of human development bringing mankind from the industrial age to the information age [2]. Now the pace of social and economic development is further accelerating, and the life rhythm is faster than before and it is also more convenient for the acquisition and storage of information, which makes the total amount of data accumulate more and more, with faster cumulative speed [3]. The Internet wave that began in the 1990s has further accelerated the accumulation of such data, and the breadth of data coverage continues to expand [4]. The data platform that the Internet provides to people is no longer the data of a place in the past [5]. This kind of information sky rocketing has not only brought people a big impact but also the imperfect network language has affected the language quality to a certain extent [6].

To this end, data mining algorithms and Bayesian networks are introduced into the evaluation of language quality training, effectively analyzing the relationship between many uncertain causal data in the process of cultivating language quality [7]. The conclusions are summarized into a general model for cultivating language quality to effectively improve language quality [8].

Following are the major novelties of this study: (1) the existing research methods for most aspects of language quality training use questionnaires to obtain data and use the analytic model to derive. This approach not only leads to an increase in research costs but also reduces the credibility of the research results [9]. However, this study is novel in introducing data mining algorithms and Bayesian networks into the assessment of language quality training, with a BOPPPS (Bridge-in, Objective/outcome, Pre-assessment, Participatory learning, Post-assessment, Summary) model designed to train the language quality of adolescents [10]. (2) In the whole process of cultivating the language quality of adolescents using the BOPPPS model, fMRI (magnetic resonance) technology is used to collect the brain wave data of the research object. The ID3 algorithm is used to mine and eliminate brainwave data that do not meet the experimental conditions, and then the remaining data are displayed in a visual way to the whole process of juvenile sentence-type inductive reasoning to obtain a more objective research result.

The organizational structure of this study is as follows: Section 2 reviews the main research status of adolescents' language quality training. In Section 3, the algorithm model is optimized. Section 4 designs a simulation experiment to verify the actual effect of the BOPPPS model, and uses the multi-level fuzzy comprehensive evaluation model to evaluate its actual effect. In Section 5, the results of the study are summarized.

2 STATE OF THE ART

Bayesian network is a tool for uncertainty reasoning, has very powerful functions, and has solved many problems that cannot be solved by conventional methods [11]. However, since the Bayesian method needs to follow the definition of each event relationship by the relevant personnel, the subsequent calculation is performed, and then the probability of occurrence of the event is calculated by other methods [12]. The definition of the event relationship is more subjective, and the calculated result also has a certain subjective tendency. Therefore, relevant research scholars added objective data to the method and combined the objective data with the definition of subjective relationships to obtain more objective calculation results [13]. After that, some scholars tried to improve the Bayesian network completely through objective observation data, without relying on subjective components of event relationship definition [14]. This raised an NP problem for scholars and this problem had also attracted the attention of many scholars, with more and more scholars joining this research area [15]. Thus, in the late 1960s, some scholars proposed a tree network model algorithm that can decompose a given probability distribution P [16]. Its implementation principle was to use cross-entropy to judge the model, but also the probability distribution P was combined to achieve the independent detection effect between variables. By this method, the best network structure was calculated [17]. The design thinking of this theory is also the basis of all subsequent research. By the end of the 1980s, researchers had improved this algorithm and extended its original application to multi-tree structures [18]. More ideas were put forward for the direct determination of the directed edge, and the conditional independence relationship was used to judge it. At this time, a more comprehensive study of the construction of directed graphs began. In the same period, some scholars proposed to use the independent relationship between variables to analyze the structural relationship between their respective nodes, and to achieve the construction of directed graphs [19]. This study has expanded the network structure into a richer research direction, not only through tree research, but more and more network structure research directions are beginning to appear. In the early 1990s, researchers first proposed the K2 algorithm, a very valuable algorithm for incorporating a priori information into Bayesian network learning and a very important algorithm in Bayesian network research work. The algorithm can be used as a priori information under the condition of knowing the order of nodes. Combined with Bayesian network, the accuracy of model and data results can be better judged. After adjusting the network and using the greedy search method, the best network structure can be gotten. As long as they meet the above methods for finding the best network structure thinking, the academic community has summarized them into structural learning methods based on evaluation and search [20]. In the same period, other scholars proposed the Kutato algorithm, and its design thinking is consistent with K2. However, it is a measure of the evaluation index, so the calculation process of the entire algorithm is more complicated than the K2 algorithm, so fewer people are using it to calculate. At the same time, some scholars calculated the best network structure based on the minimum description length, only using search and evaluation. This method also does not require prior information. Later, some scholars proposed a new algorithm for evaluating search by the minimum information length. It can even calculate the optimal network structure without knowing the order of the nodes. Many excellent algorithms have been designed for the idea of completing structural learning through independent testing. In the early 1990s, some scholars improved the Boundary DAG algorithm and realized the construction of the network structure only when a part of the node order and prior information were known.

Rydenvald [21] studied the use of languages and attitudes among adolescents in some European schools, with particular attention to the interrelationship between the third culture and elite bilingualism among adolescents, showing that there is a certain correlation between teaching languages and language education attitudes in different international education curricula. In the traditional sense, the definition of language ability is the preparation state and ability of individuals to learn the language, which is long-term and stable expression ability. It can be seen from the existing studies that working memory is also an important part of language ability [22]. Li used the meta-analysis analysis method to study the effect of language ability in second language learning. The results show that language ability, as the cognitive ability of young people, is a very important condition in the learning process of the second language [23]. After the gradual improvement, the Bayesian network is used in various fields due to its powerful performance and its excellent computing power in uncertain causal relationships. The evaluation of the effect of adolescents' language quality training is also one of the most suitable directions for its application. Bayesian network is a new technology, the application research on the evaluation of adolescents' language quality training is limited, the data parameters are not complete, and the specific training scheme is not perfect, which has greatly limited the development of Bayesian network capabilities and the improvement of the language quality of adolescents.

Therefore, in view of the shortcomings of the existing research, this study uses the data mining ID3 algorithm to improve the data parameters of Bayesian network in the training of adolescent language quality. That is to say, the data mining algorithm is used to mine the data clusters collected in the process of adolescents' language quality training, and then the value data parameters hidden in the data cluster are obtained. Then these data parameters are applied to the Bayesian network to effectively improve the cultivation ability of the Bayesian network, improving the BOPPPS model designed in this study to train the language quality of adolescents.

3 METHODOLOGY

It is known from the above description that fMRI (magnetic resonance) technology is used to collect brain waves of the human brain to realize real-time imaging of the activity between the brain regions to study the juvenile inductive inference ability of adolescents. But there are two major difficulties in this research method: (1) To better visualize the activity between different regions of the human brain, human brain brainwave data collected by fMRI (magnetic resonance) technology is often calculated in terabytes and the value of information existing between these massive human brain-wave data is difficult to explore conventionally; (2) The causal link between human brainwave data collected by fMRI (magnetic resonance) technology is often uncertain, which makes it difficult to study the brainwave data of these people.

In response to the above problems and comprehensive research results, data mining algorithm and Bayesian network design are used to achieve a BOPPPS model in this study with the BOPPPS model used for adolescents' language quality. The Bayesian network runs through the whole research process, mining and analyzing the fMRI data through the Bayesian network, and thus proposing an improved BOPPPS model using Bayesian network, scientific and reasonable quantitative analysis can be effectively realized, while avoiding the influence of invalid data in the fMRI and providing the most scientific basis for the subsequent BOPPPS model. The BOPPPS model mainly uses the ID3 data mining algorithm to solve the problem that massive human brain wave data collected by fMRI (magnetic resonance) technology is difficult to mine using conventional methods. At the same time, Bayesian network is used to solve the problem of uncertain causal connection between human brain wave data collected by fMRI (magnetic resonance) technology.

3.1 ID3 Algorithm

The ID3 algorithm is a criterion for judging the selected content based on the degree of attenuation exhibited by the information entropy. ID3 is developed by continuous research and analysis on two types of classification problems and is transformed into mathematical expressions as:

Let $E = F_1 * F_2 * \ldots * F_n$ be the n-dimensional finite vector space, where F_j is the finite discrete symbol set. The element in E $e = \langle V_1, V_2, ..., V_n \rangle$ is an instance, among them $V_j \in F_j$, $j = 1, 2, \ldots, n$. Let P and N be two sets of instances of E and F, which are called positive and negative sets respectively as shown in Fig. 1.

The amount of information required to determine is expressed by the following formula:

$$E(E) = -\frac{P_i}{P_i + N_i} \log \frac{P_i}{P_i + N_i} - \frac{N_i}{P_i + N_i} \log \frac{N_i}{P_i + N_i} \quad (1)$$

If attribute A is used as the root of the decision tree, A has V values V_1, V_2, …, V_v which divides E into V subsets. Suppose E_i contains P_i positive examples and N_i counter examples. The information entropy of subset E_i is $E(E_i)$.

$$E(E_i) = \frac{P_i}{P_i + N_i} \log \frac{P_i}{P_i + N_i} + \frac{N_i}{P_i + N_i} \log \frac{N_i}{P_i + N_i} \quad (2)$$

The information entropy after the attribute A as the root is $E(A)$:

$$E(A) = \sum_{r=1}^{v} \frac{P_i + N_i}{P + N} E(E_i) \quad (3)$$

Therefore, the information gain $I(A)$ with the attribute as the root is:

$$I(A) = E(E) - E(A) \quad (4)$$

ID3 selects the attribute A that makes $I(A)$ maximum ($E(A)$ minimum) as the root node. Let the sample set S have a total of C samples, and the number of samples per class is $P_i = (i = 1, 2, 3, \ldots, e)$. If the attribute A is the lowest level of the decision tree, with V values V_1, V_2, …, V_v, it divides E into V subsets $[E_1, E_2, \ldots, E_v]$. Assuming that the number of samples containing j in E_i is $P_{ij} = 1, 2, \ldots, c$, then the information amount of subset E_i is $E(E_i)$:

$$E(E_i) = -\sum_{j=1}^{c} \frac{P_{ij}}{|E_i|} \log \frac{P_{ij}}{E_i} \quad (5)$$

The information entropy classified by A is:

$$E(A) = \sum_{i=1}^{v} \frac{|E_i|}{E} E(E_i) \quad (6)$$

Selecting attribute A minimizes $E(A)$ in Eq. (6), and the information gain becomes more.

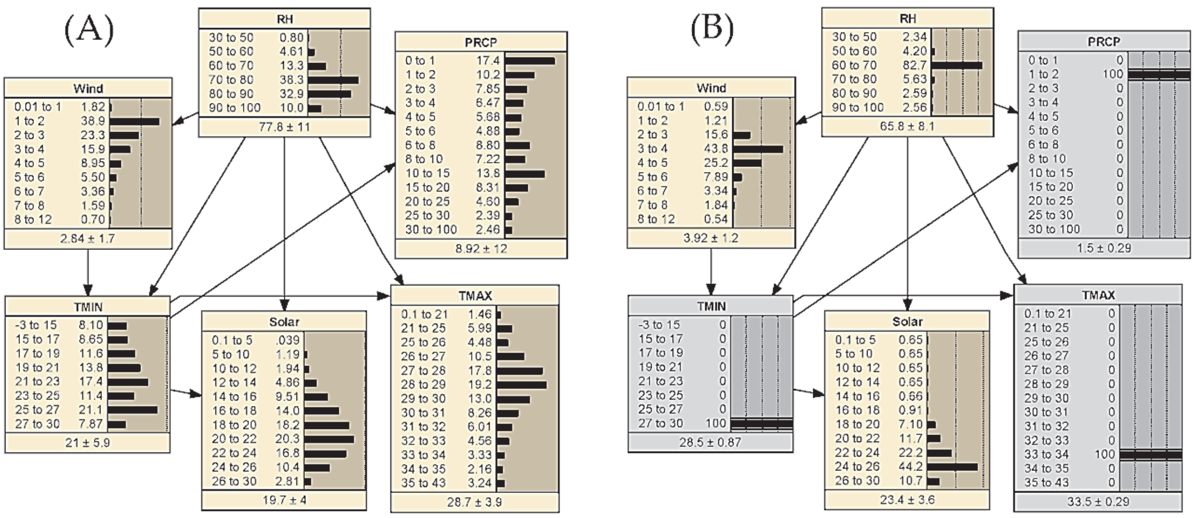

Figure 1 Positive and negative examples

3.2 Bayesian Network

A Bayesian network is a common data model used to predict the causal link between different events. The basic principle of the model is mainly to calculate the causality between different data by using DAG (Directed Acyclic Graph) as shown in Fig. 2. The principle is as follows:

The raw data needs to be processed first to generate an ordered sequence of data.

If most of the level ratios fall within the allowable coverage area $(e^{\frac{-2}{n+1}}, e^{\frac{2}{n+1}})$, then the Bayesian network model can be established and the prediction function can be implemented. Otherwise, properly preprocess the data. The processing method is as follows:

$$X^{(0)}(t) = \frac{X^{(0)}(t-1) + 2X^{(0)}(t) + X^{(0)}(t+1)}{4} \quad (7)$$

$$X^{(0)}(n)=\frac{X^{(0)}(n-1)+3X^{(0)}(n)}{4} \tag{8}$$

$$X^{(0)}(n)=\frac{X^{(0)}(n-1)+3X^{(0)}(n)}{4} \tag{9}$$

After processing, $X^{(1)}(k)=\sum_{n=1}^{k}X^{(0)}(n)$, a new series is gotten.

$$X^{(1)}(k)=\sum_{n=1}^{k}X^{(0)}(n) \tag{10}$$

The differential equation for this series is as follows:

$$\frac{\mathrm{d}X^{(1)}}{\mathrm{d}t}+aX^{(1)}=u \tag{11}$$

a in the formula is the development gray number; u is the endogenous control gray number.

Let $Y_n=[X^{(0)}(2),X^{(0)}(3),\ldots,X^{(0)}(n)]^{\mathrm{T}}$ and $\hat{\alpha}$ be the parameter vector $\hat{\alpha}=\begin{pmatrix}a\\u\end{pmatrix}$ to be estimated.

$$B=\begin{bmatrix}-\frac{1}{2}(X^{(1)}(1)+X^{(1)}(2)) & 1\\ -\frac{1}{2}(X^{(1)}(2)+X^{(1)}(3)) & 1\\ \cdots & \cdots\\ -\frac{1}{2}(X^{(1)}(n-1)+X^{(1)}(n)) & 1\end{bmatrix} \tag{12}$$

The model can then be expressed as $Y_n=B\hat{\alpha}$ obtained by the least squares method: $\hat{\alpha}=\left(B^TB\right)^{-1}B^TY_n$ solves the differential equation, then the predicted discrete time response function:

$$\hat{X}^{(1)}(t+1)=\left[X^{(0)}(1)-\frac{u}{a}\right]e^{-at}+\frac{u}{a},t=0,1,2\ldots,n-1 \tag{13}$$

$\hat{X}^{(1)}(t+1)$ is the accumulated predicted value, and the predicted value is restored as:

$$\hat{X}^{(0)}(t+1)=\hat{X}^{(1)}(t+1)-\hat{X}^{(1)}(t) \tag{14}$$

The original data sequence with variable $x^{(0)}$:

$$x^{(0)}=\left\{x^{(0)}(1),x^{(0)}(2),\ldots.x^{(0)}(n)\right\} \tag{15}$$

The first generation cumulative generation module $x^{(1)}$ is generated by the accumulation generation algorithm:

$$x^{(1)}=\left\{x^{(1)}(1),x^{(1)}(2),\ldots.x^{(1)}(n)\right\} \tag{16}$$

$x^{(1)}(k)=x^{(1)}\sum_{i=1}^{k}x^{(0)}(i)$ in the formula.

A differential equation consisting of a first-order gray module $x^{(1)}$:

$$\frac{\mathrm{d}x^{(1)}}{\mathrm{d}t}+ax^{(1)}=b \tag{17}$$

After simplification can get:

$$\frac{\mathrm{d}x}{\mathrm{d}t}=\lim_{\Delta t\to 0}\frac{x(t+\Delta t)-x(t)}{\Delta t} \tag{18}$$

If expressed in discrete form, as shown in Fig. 3, the derivative term can be written as:

$$\begin{aligned}\frac{\Delta x}{\Delta t}&=\frac{x(k+1)-x(k)}{k+1-k}=x(k+1)-x(k)\\&=a^{(1)}[x(k+1)]\end{aligned} \tag{19}$$

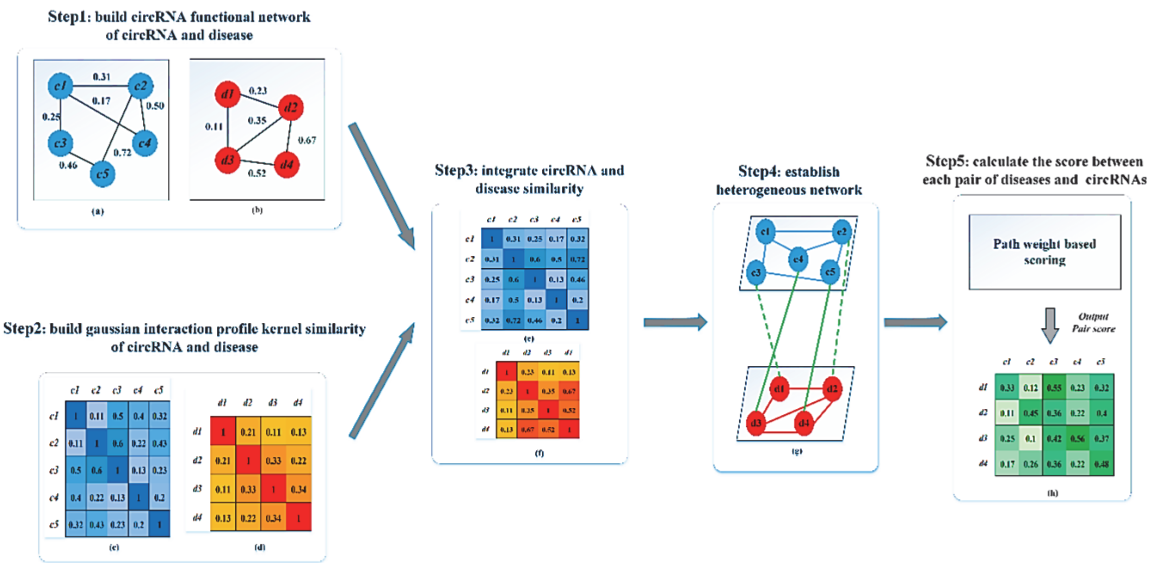

Figure 2 Randomness versus raw data

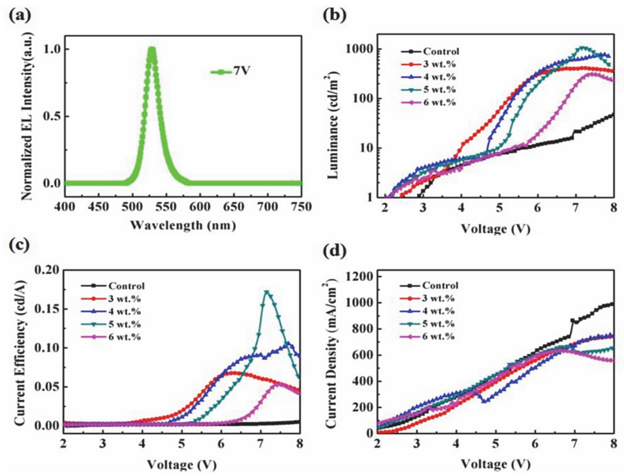

Figure 3 After pre-processing, the data is processed by a cumulative generation process

In this case, the x value can only take the average of the time k and $k+1$, namely: $\frac{1}{2}[x(k+1)+x(k)]$. The differential equation can be rewritten as:

$$
\begin{aligned}
& a^{(1)}[x^{(1)}(k+1)]+\frac{1}{2}a[x^{(1)}(k+1)+x^{(1)}(k)]=b \\
& k=1, x^{(0)}(2)+\frac{1}{2}a[x^{(1)}(1)+x^{(1)}(2)]=b \\
& k=2, x^{(0)}(3)+\frac{1}{2}a[x^{(1)}(2)+x^{(1)}(3)]=b \\
& \vdots \\
& k=N-1, x^{(0)}(N)+\frac{1}{2}a[x^{(1)}(N-1)+x^{(1)}(N)]=b
\end{aligned}
\tag{20}
$$

Written in matrix form, there are:

$$
\begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(N) \end{bmatrix}
=
\begin{bmatrix}
-\frac{1}{2}[x^{(1)}(1)+x^{(1)}(2)] & 1 \\
-\frac{1}{2}[x^{(1)}(2)+x^{(1)}(3)] & 1 \\
\vdots & \vdots \\
-\frac{1}{2}[x^{(1)}(N-1)+x^{(1)}(N)] & 1
\end{bmatrix}
\begin{bmatrix} a \\ b \end{bmatrix}
\tag{21}
$$

In the above equations, Y and X are known quantities, and, B is a pending parameter. Since the variables are only a and b,

But the least squares solution can be obtained by the least squares method. So the above equation can be directly rewritten as:

$$
Y = X\hat{B} + E \tag{22}
$$

The variable E in this formula is the error term.

$$
\min\left\|Y - X\hat{B}\right\|^2 = \min\left(Y - X\hat{B}\right)^{\left(Y - X\hat{B}\right)} \tag{23}
$$

Using the matrix derivation formula can get:

$$
\hat{B} = (X^T \mathrm{Y}) = \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} \tag{24}
$$

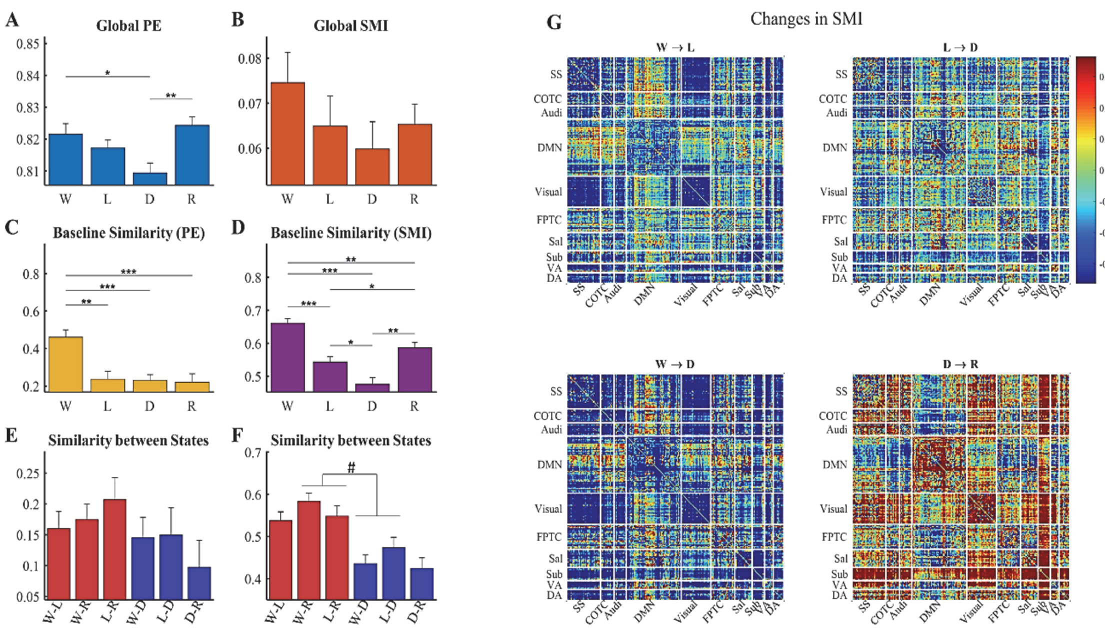

Figure 4 The smaller the variance, the higher the prediction accuracy

Going back to the original differential equation, there are:

$$
\frac{\mathrm{d}x^{(1)}}{\mathrm{d}t} + \hat{a}\,x^{(1)} = \hat{b} \tag{25}
$$

Solving the above equation can get

$$
\begin{aligned}
& x^{(1)}(t) = [x^{(1)}(0) - \frac{b}{a}]e^{-at} + \frac{b}{a} \\
& x^{(1)}(k+1) = [x^{(0)}(0) - \frac{b}{a}]e^{-at} + \frac{b}{a}
\end{aligned}
\tag{26}
$$

To simplify the calculation, the above formula can be simplified to the following formula:

$$S_e^{\,2} = \frac{1}{N}\sum_{k=1}^{N}\left[e^{(0)}(k) - e\right]^2 \tag{27}$$

$$S_x^{\,2} = \frac{1}{N}\sum_{k=1}^{N}\left[x^{(0)}(k) - x^{(0)}\right]^2 \tag{28}$$

It can be seen from Fig. 4 that the smaller the square difference calculated based on the aforementioned formula, the higher the prediction accuracy obtained

4 RESULTS ANALYSIS

To test the practical effect of the BOPPPS model designed by data mining algorithm and Bayesian network design in the process of cultivating adolescents' language quality, simulation experiments on the BOPPPS model and evaluate its actual effect are carried out through a multi-level fuzzy comprehensive evaluation model.

4.1 Experimental Environment

According to the amount of data and the need for calculation, this study needs to build an operation platform; according to the number of research objects, we need the corresponding experimental platform and the main server node. To simplify the experiment difficulty, four PCs are used to build a simulation platform. One PC acts as the main server node, and three PCs are connected as Ethernet as the background server as the experimental platform of the BOPPPS model.

4.2 Experimental Steps

A total of 642 adolescents were selected to inform of the contents and purpose of the study before conducting the study. The BOPPPS model cultivates adolescents' language quality mainly in accordance with the following six stages: (1) Introducing (B), the process is mainly to introduce a cognitive premise to the adolescents who are the research subjects; (2) Goal (O), the process is mainly to set an expected target for these adolescents; (3) Pretest (P), a pre-test should be conducted for them before conducting the study; (4) Participatory learning (P), which is mainly to let them actively participate in the activities of language quality training; (5) Post-test (P), after the completion of the study, a post-test is required for them; (6) Summarize (S), this process is mainly to evaluate the actual effect of adolescents' language quality training as shown in Fig. 5.

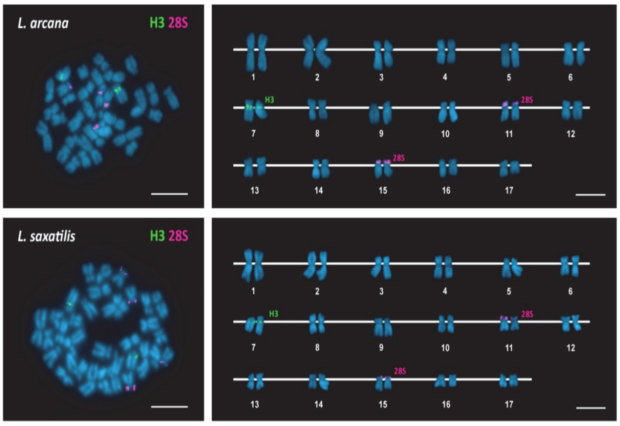

Figure 5 To evaluate the practical effect of language quality training for young people

In the whole process of cultivating adolescents' language quality by using the BOPPPS model, fMRI (magnetic resonance) technology is used to collect brain wave data of research subjects, using the ID3 algorithm to mine and eliminate brainwave data that do not meet the experimental conditions. Then the remaining data is displayed in a visual way to the whole process of adolescent sentence-type inductive reasoning to ensure that an objective scientific research result can be obtained.

4.3 Experimental Evaluation

After obtaining a series of visual images of brainwave data, it is necessary to use the multi-level fuzzy comprehensive evaluation model to analyze and evaluate these images to evaluate the actual training effect of the BOPPPS model on adolescent language quality.

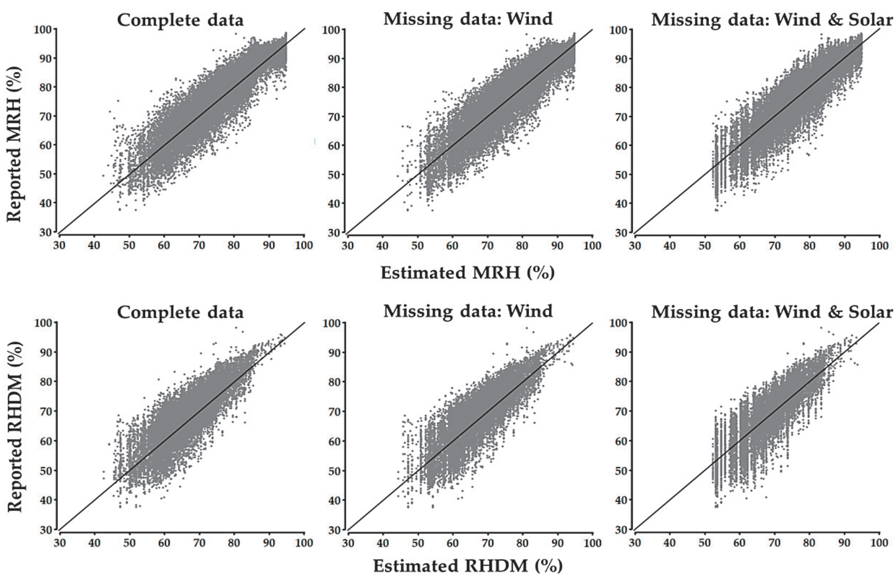

Figure 6 The relative weight of the criterion layer index in the criterion layer

Figure 7 Using data mining algorithm and Bayesian network

The multi-level fuzzy comprehensive evaluation model is based on cognitive science and fuzzy mathematics as a theoretical basis and is created by an evaluation method as shown in Fig. 6. Fig. 6 reflects the relative weight value corresponding to the criterion layer index in the residential layer. First, establish a hierarchical evaluation matrix $\boldsymbol{R}$, and then calculate the product M_i of each row element of the judgment matrix $\boldsymbol{R}$. The target metric layer weight vector is obtained by calculating the n root square $\overline{w_i}$ of M_i and normalizing it.

Finally, the index weights of different weight vectors are set, with the data information recorded in the experiment process substituted into the above equation, and the evaluation result can be obtained as shown in Fig. 7. The data in the figure shows that the BOPPPS model proposed by the data mining algorithm and Bayesian network design proposed in this study plays a positive role in improving the language quality of adolescents. This study is to use fMRI nuclear magnetic resonance technology to obtain brain wave data of teenagers, and after the screening, it is analyzed by big data technology. The language quality of teenagers is a comprehensive ability. Bayesian network can analyze the effect of language quality training on teenagers. Using the BOPPPS model can directly use the data obtained by fMRI nuclear magnetic resonance technology so that conceptual language quality ability can be transformed into a concept that can be quantitatively analyzed. Through this model, we can directly analyze the differences in language quality of teenagers, and thus can carry out targeted training.

5 CONCLUSION

To test the practical effect of the BOPPPS model designed by using data mining algorithm and Bayesian network in the training of language quality a total of 642 adolescents were selected to carry out specific experiments, with 1 and 3 PC machines as the main server nodes and experimental. The fMRI model is proposed by visualizing the language quality through BOPPPS NMR. The following conclusions can be drawn. The BOPPPS model based on data mining algorithm and Bayesian network can effectively improve the language quality of adolescents. This lifting effect is mainly reflected in the mining and analysis of fMRI brain wave data based on Bayesian network. The resulting BOPPPS model and the traditional simple proposed BOPPPS are more scientific and targeted. During the training of adolescents' language quality, the proposed BOPPPS model can be used directly for teaching applications, which can get a better application effect.

There are some limitations in this study, such as small sample size, short experimental time, lack of follow-up, etc. in the later stage, it is necessary to expand the sample size of the study, as long as possible, follow up the speech improvement in the later stage, and further study different types of subjects.

Acknowledgements

This study was supported by Zhejiang Province 2016 visiting scholar professional development project (117).

6 REFERENCES

[1] Vela-Oxolon. I. A., Requejo-Mirez, M. S., Cubillas-Cochachin, C. G., Perez-Mantari, L. K., & Alfaro-Paredes, E. A. (2019). Analyzing the classification of technical standards for the management of information technology infrastructure and services. *DYNA*, *94*(5), 484-484. https://doi.org/10.6036/9303

[2] Urrutia-Azcona, K., Stendorf-Sørensen, S., Molina-Costa, P., & Flores-Abascal, I. (2019). Smart Zero Carbon City: key factors towards smart urban decarbonisation. *DYNA*, *94*(6), 676-683. https://doi.org/10.6036/9273

[3] Caldwell, C. E. (2019). Adolescent brain development and gender: predictors of future reading habits. *Journal of Neurology, Neurosurgery & Psychiatry*, *90*(2), 235-237. https://doi.org/10.1136/jnnp-2018-318094

[4] Yu, G. H., Wang, L. Q., Wan, S. C., Qi, L. Q., & Xu, Y. W. (2021). Tensor factorization with total variation for Internet traffic data imputation. *Pacific journal of optimization*, *17*(3), 486-505.

[5] Silva, M. N., Naspitz, C., & Solé, D. (2001). Evaluation of quality of life in children and teenagers with allergic rhinitis: adaptation and validation of the Rhinoconjunctivitis Quality of Life Questionnaire (RQLQ). *Allergologia et Immunopathologia*, *29*(4), 111-118. https://doi.org/10.1016/S0301-0546(01)79042-8

[6] Skrynnik, A., Staroverov, A., Aitygulov, E., Aksenov, K., Davydov, V., & Panov, A. I. (2021). Hierarchical Deep Q-Network from imperfect demonstrations in Minecraft. *Cognitive Systems Research*, *65*(7676), 74-78. https://doi.org/10.1016/j.cogsys.2020.08.012

[7] Kazanidis, I., Valsamidis, S., Gounopoulos, E., & Kontogiannis, S. (2020). Proposed S-Algo+data mining algorithm for web platforms course content and usage evaluation. *Soft Computing*, *24*(19), 14861-14883. https://doi.org/10.1007/s00500-020-04841-8

[8] Feng, W., Zhao, Y., Zhao, Z., Hong, J., Luo, Z., & Yin, L. (2015). Study on the use of big data to promote food and drug smarter supervision. *Journal of Food Safety and Quality*, *6*(1), 354-360.

[9] Nemèšanu, F. & Pãžnzaru, F. (2017). Smart city management based on IoT. *Smart Cities and Regional Development (SCRD) Journal*, *1*(1), 91-97. https://doi.org/10.25019/scrd.v1i1.12

[10] Bellini, P., Nesi, P., & Pantaleo, G. (2022). IoT-enabled smart cities: A review of concepts, frameworks and key technologies. *Applied Sciences*, *12*(3), 1607. https://doi.org/10.3390/app12031607

[11] Sánchez, V., Muñoz-Fernández, N., & Ortega-Ruíz, R. (2015). Cyberdating Q_A: An instrument to assess the quality of adolescent dating relationships in social networks. *Computers in Human Behavior*, *48*, 78-86. https://doi.org/10.1016/j.chb.2015.01.006

[12] Du, C., Zhu, H., & Fang, Y. (2015). Institutional complementarities and differences in pharmaceutical circulation systems between foreign countries and China. *Finance & Trade Economic*, *4*, 109-120.

[13] Guo, Q. (2015). Application of Cloud Computing and IOT to Smart City Construction. *Journal of Chongqing University of Science and Technology: Natural Science Edition*, *17*(3), 95-97.

[14] Chen, X. (2019). The development trend and practical innovation of smart cities under the integration of new technologies. *Frontiers of Engineering Management*, *6*(4), 485-502. https://doi.org/10.1007/s42524-019-0057-9

[15] Fuligni, A. J., Tsai, K. M., Krull, J. L., & Gonzales, N. A. (2015). Daily concordance between parent and adolescent sleep habits. *Journal of Adolescent Health*, *56*(2), 244-250. https://doi.org/10.1016/j.jadohealth.2014.09.013

[16] Patil, J. S. & Sarasija, S. (2012). Pulmonary drug delivery strategies: A concise, systematic review. *Lung India: Official Organ of Indian Chest Society*, *29*(1), 44-49. https://doi.org/10.4103/0970-2113.92361

[17] Khanna, A. & Kaur, S. (2020). Internet of things (IoT), applications and challenges: a comprehensive review. *Wireless Personal Communications*, *114*(2), 1687-1762. https://doi.org/10.1007/s11277-020-07446-4

[18] Gómez Romero, C. D., Díaz Barriga, J. K., & Rodríguez Molano, J. I. (2016, June). Big data meaning in the architecture of IoT for smart cities. *International Conference on Data Mining and Big Data*, 457-465. https://doi.org/10.1007/978-3-319-40973-3_46 8

[19] Kim, T. H., Ramos, C., & Mohammed, S. (2017). Smart city and IoT. *Future Generation Computer Systems*, *76*, 159-162. https://doi.org/10.1016/j.future.2017.03.034

[20] Cvar, N., Trilar, J., Kos, A., Volk, M., & Stojmenova Duh, E. (2020). The use of IoT technology in smart cities and smart villages: similarities, differences, and future prospects. *Sensors*, *20*(14), 3897. https://doi.org/10.3390/s20143897

[21] Rydenvald, M. (2015). Elite bilingualism? Language use among multilingual teenagers of Swedish background in European Schools and international schools in Europe. *Journal of Research in International Education*, *14*(3), 213-227. https://doi.org/10.1177/1475240915614935

[22] Singleton, D. (2017). Language aptitude: Desirable trait or acquirable attribute? *Studies in Second Language Learning and Teaching*, *7*(1), 89-103. https://doi.org/10.14746/ssllt.2017.7.1.5

[23] Li, S. (2015). The associations between language aptitude and second language grammar acquisition: A meta-analytic review of five decades of research. *Applied Linguistics*, *36*(3), 385-408. https://doi.org/10.1093/applin/amu054

Contact information:

Jie CHEN, Master, Associate Professor
Yiwu Industrial & Commercial College,
Jinhua, Zhejiang, China
E-mail: 15057806960@ywicc.edu.cn