

Ideological and political assessment system based on data visualization

Ning Wang¹

¹Xi'an Aeronautical Polytechnic Institute, Shanxi, Xian, 710089, China.

Abstract. Data visualization is a very important research topic of current computer science and technology. It is a theoretical method and technology that uses the technology of computer processing massive and complex data to transform the data results produced in daily life and production into graphics or images to show to the information receiver, and carry out human-computer interaction. Data visualization is a research hotspot in the field of Computer Science in recent years. Data visualization technology is mainly used in Applied Mathematics. With the continuous development of data visualization technology, the application of data visualization technology will be widely used in daily life. Higher vocational colleges, as educational institutions to cultivate high-level skilled and application-oriented talents, should strengthen the foundation, broaden the occupation, improve the specialty, and strengthen the ideological and political construction. Therefore, the ideological and Political Theory Teachers in higher vocational colleges should practice the education policies of the party and the state, ensure the correct direction of Ideological and political education, and actively explore and re select the ideological and political theory line through a variety of evaluation methods. With the continuous development and replacement of high and new technology, the teaching of higher vocational colleges should also keep up with the times. Data visualization technology should also be integrated into the ideological and political assessment of students. This paper introduces the data visualization technology and methods, puts forward some key and difficult points and breakthrough points in the application of data visualization in the ideological and political assessment system, and introduces several mathematical model algorithm evaluation examples, hoping to help the data visualization of the ideological and political assessment system in higher vocational colleges.

1 Introduction

It is a key educational project for every college to build high-quality ideological and political norms for students and strive to cultivate socialist builders and successors with Chinese characteristics. Through scientific education, we can realize the strategy of strengthening the country with talents, realize the comprehensive construction of a well-off society, and accelerate the advancement of socialist modernization. It has great and extensive significance. The arrival of the new socialist era indicates that China has entered a new stage of personnel training. Students in Higher Vocational Colleges learn all kinds of skills. As applied talents of social skills, they are distributed in all posts of the society and should set an example [2]. At present, there are some problems in the assessment of Ideological and political course in most colleges, such as single form, few levels of assessment standards, and not rich assessment contents [3]. In order to better carry out the ideological and political assessment, in view of the common problems in the assessment of Ideological and political course, this paper puts forward the improvement of the assessment system of Ideological and political theory course in Higher Vocational Colleges

from the perspectives of assessment form, assessment standard and assessment content, so as to make it a diversified and modern assessment method [4].

In addition to the necessary cultural courses, the main part of the students in higher vocational colleges is the skill learning of the corresponding majors [5]. The professor of professional skills corresponding professional teachers. The ideological and political assessment of teachers is also a very important part, even if they carry out the ideological and political assessment when carrying out the relevant teacher assessment, but the ideological and political examination is not a one-time assessment. Ideological and political assessment needs to be continuous and repeated [6]. Teachers, as the main personnel who teach students' professional skills and ideological and political education in higher vocational colleges, their ideological and political assessment is particularly important. Therefore, this paper proposes that the previous ideological and political assessment system only focuses on the ideological and political assessment of students, and ignores the ideological and political assessment of teachers. In the system of Ideological and political assessment, teachers' Ideological and political assessment should be added. Only by establishing the correctness of Ideological and political education from the

root of teaching can we eliminate the factor of professors who have problems in students' Ideological and political assessment, which is very necessary [7].

In the current era of rapid development of computer technology, higher vocational colleges should also keep up with the pace of the times. Higher vocational colleges should add data visualization technology to the ideological and political assessment system designed by various departments. Data visualization technology is based on the processing and analysis of a large number of data by computer, which makes the complex and disordered data vivid and become direct and simple visual information such as tables, pictures and videos [9]. The use of data visualization system can reduce the complexity and workload of managers. Improve the speed of readers' direct reading information, and have the best of both worlds. Therefore, the application of data visualization in the ideological and political assessment system of higher vocational colleges is urgent [10].

2 Algorithm establishment

2.1 iterative algorithm

Most of the data of this system can be collected by the collector or imported by human, but some data can only be obtained by calculation, such as the average value of speed, Cpk, variance and other data. These data must be obtained by calculating the original speed value by computer. The calculation formula of variance is as follows:

$$v_n = \frac{1}{n} \sum_{i=1}^n (x_i - r)^2 \quad (1)$$

Where x_i is the value of each item and r is the average value of all items. The variance of a batch of data can be easily calculated by formula (4-1). However, the space complexity and time complexity of the algorithm are $O(n)$ and $O(n)$. Because the original data collected are increasing and the increasing speed is very fast, the space and time requirements of the algorithm will increase with the increase of time, If it is not restricted, it will even cause memory overflow and other problems, so we must optimize the algorithm to reduce the consumption of space and time. In this paper, the iterative algorithm is used to optimize the variance algorithm. The space complexity of the algorithm is reduced to $O(1)$, and the time complexity is also reduced to $O(1)$

The general formula for calculating the mean value is as follows:

$$A_n = \frac{1}{n} \sum_{i=1}^n x_i \quad (2)$$

By disassembling formula (2), the recursive formula of mean value is obtained

$$A_n = \frac{1}{n} [(n-1)A_{n-1} + X_n] = A_{n-1} + \frac{x_n - A_{n-1}}{n} \quad (3)$$

By introducing the mean recurrence formula (3) into variance (4), we can get the following results:

$$V_n = \frac{1}{n} \sum_{i=1}^n \left[x_i - \left(A_{n-1} + \frac{x_n - A_{n-1}}{n} \right) \right]^2 \quad (4)$$

Recombining formula (4) to get:

$$V_n = \frac{1}{n} \sum_{i=1}^n \left[(x_i - A_{n-1}) - \frac{x_n - A_{n-1}}{n} \right]^2 \quad (5)$$

Expand formula (5)

$$V_n = \frac{1}{n} \sum_{i=1}^n \left[(x_i - A_{n-1})^2 + \left(\frac{x_n - A_{n-1}}{n} \right)^2 - 2(x_i - A_{n-1}) \left(\frac{x_n - A_{n-1}}{n} \right) \right] \quad (6)$$

The constant is extracted from formula (6)

$$V_n = \left(\frac{x_n - A_{n-1}}{n} \right)^2 + \frac{1}{n} \sum_{i=1}^n \left[(x_i - A_{n-1})^2 - 2(x_i - A_{n-1}) \left(\frac{x_n - A_{n-1}}{n} \right) \right] \quad (7)$$

Because:

$$\sum_{i=1}^n x_i = \sum_{i=1}^{n-1} x_i + x_n \quad (8)$$

So we can take x_i in formula (7) as an example_ I extract, make

$$S_n = x_n - A_{n-1} \quad (9)$$

$$S_i = x_i - A_{n-1} \quad (10)$$

The results are as follows

$$V_n = \left(\frac{S_n}{n} \right)^2 + \frac{1}{n} \left[S_n^2 - \frac{2S_n^2}{n} \right] + \frac{1}{n} \sum_{i=1}^{n-1} \left[S_i^2 - 2S_i \frac{S_n}{n} \right] \quad (11)$$

Formula (11) is combined with constant terms

$$V_n = \frac{n-1}{n^2} S_n^2 + \frac{1}{n} \sum_{i=1}^{n-1} S_i^2 - \frac{2S_n}{n^2} \sum_{i=1}^{n-1} S_i \quad (12)$$

$$\sum_{i=1}^{n-1} S_i = \sum_{i=1}^{n-1} x_i - \sum_{i=1}^{n-1} A_{n-1} = 0 \quad (13)$$

Finally, the formula (12) is simplified

$$V_n = \frac{n-1}{n^2} S_n^2 + \frac{1}{n} \sum_{i=1}^{n-1} S_i^2 \quad (14)$$

Because:

$$V_n = \frac{n-1}{n^2} S_n^2 + \frac{1}{n} \sum_{i=1}^{n-1} S_i^2 \quad (15)$$

Take formula (15) into (14)

$$V_n = \frac{n-1}{n^2} S_n^2 + \frac{n-1}{n} V_{n-1} \quad (16)$$

Finally, formula (9) is brought back to formula (16)

$$V_n = \frac{n-1}{n^2} (x_n - A_{n-1})^2 + \frac{n-1}{n} V_{n-1} \quad (17)$$

According to formula (3) and formula (17), the variance can be calculated by iterative recursive algorithm. At this time, the space complexity and time complexity are reduced to $O(1)$, which greatly improves the efficiency of the program.

2.2 k-means clustering algorithm

K-means is a clustering method based on partition. It divides samples according to their similarity. Similarity is generally measured by Euclidean distance.

Suppose that there are n samples in the data set s , and the Euclidean distance between any two samples X and Y in S is

$$\text{dist}(X, Y) = (x_i y_i)^2 \quad (18)$$

Suppose that the k -Subsets after clustering are (c_1, c_2, \dots, c_k) , and the objective convergence function of K-means is:

$$E = \sum_{i=1}^k \sum_{x \in c_i} |x - \mu_i| \quad (19)$$

The algorithm flow is as follows:

(1) Input the value of the target cluster K , and randomly select k samples as the initial cluster centroid.

(2) After traversing the remaining samples, the similarity between each sample and each centroid is calculated, and the samples are divided into the most similar clusters.

(3) The clustering centroid is recalculated, and the sample mean value of each cluster is used as the new clustering centroid.

(4) Repeat steps (2) and (3) until the algorithm converges.

(5) Output the final cluster centroid and K cluster partition.

The principle of K-means algorithm is simple and easy to implement. When the feature difference is obvious, it can usually get better clustering effect. However, the selection of K value and initial cluster centroid is not referential, so a lot of experiments are needed to determine the optimal pear setting. The initial centroid is randomly selected, so it is easy to select outliers or make the clustering converge locally. When the outliers are selected, a small amount of noise will cause great interference to the average value, which greatly reduces the clustering effect.

2.3 k-means algorithm optimization

K-means algorithm is a partition based algorithm, which has an obvious disadvantage. The initial centroid is randomly selected according to the K value. In this way, it is easy to select the isolated point or fall into the local optimum, which often requires a lot of experiments to select the optimal result. At present, most of the researchers improve the algorithm by density optimization or distance optimization, but these improvements only consider the interval of the initial centroid to avoid falling into local optimization, but do not consider that the selection of outliers will also affect the clustering effect. In this paper, a method based on minimum variance and similarity optimization is proposed to improve the k-means algorithm, optimize the selection of initial centroid, avoid falling into the local optimal solution at the same time, avoid selecting isolated points. The minimum variance can ensure that the initial centroids are in the sample dense area, and the combination of similarity threshold can ensure that the initial centroids are isolated.

3 Establishment of ARIMA time series prediction model

Time series prediction is to analyze a series of factors reflected in the time series and deduce the changes in the future time. ARIMA (P, D, q) model is called autoregressive integral moving average model.

Autoregressive process: mainly through the historical observation value and the current interference value of linear combination prediction. If a linear process excluding mean and deterministic components can be expressed as:

$$y_t = \sum_{i=1}^p \varphi_i y_{t-i} + \varphi_i \quad (20)$$

Where t is the time independent variable, y is the observed value, and φ_i is the time independent variable. I is the autoregressive coefficient, φ_i is a white noise process, P is the order of autoregression, which is called y_t is an autoregressive process of order P, expressed by AR (P).

Moving average process: mainly through the historical interference value and the current interference value of linear combination prediction. If a linear process excluding mean and deterministic components can be expressed as:

$$y_t = \sum_{i=1}^p \theta_i \mu_{t-i} + \mu_t \quad (21)$$

Where t is the time independent variable, y is the observed value, θ_i is the moving average coefficient, μ_t is a white noise process, q is the order of moving average, and y_t is the process is a moving average process of order q, expressed by MA (q).

Autoregressive moving average process: if a linear stochastic process excluding the mean and determining the components is composed of autoregressive process and moving average process, it is called autoregressive moving average process, which is recorded as ARMA (P, q), and the expression is: P

$$y_t = \sum_{i=1}^p \varphi_i y_{t-i} + \mu_t \sum_{i=1}^p \theta_i \mu_{t-i} \quad (22)$$

4 Evaluation results and research

Table 1 data types of Ideological and political assessment system

Data Type	Data length	Notes
VAECHAR2	20	Student number
DETA	7	Answer time
VAECHAR2	20	Fraction

According to table 1, the main data used in this study are: student number, students' answer time and score of Ideological and political course examination. These data are usually the main basis for the ideological and political assessment of students. These main data are input into the computer database through the computer user control end, and then the database is intelligently integrated and analyzed, and finally output to the computer screen of the auditor through the data visualization algorithm. Through the processing of data visualization system, the ideological and political assessment results of higher vocational students are clear at a glance.

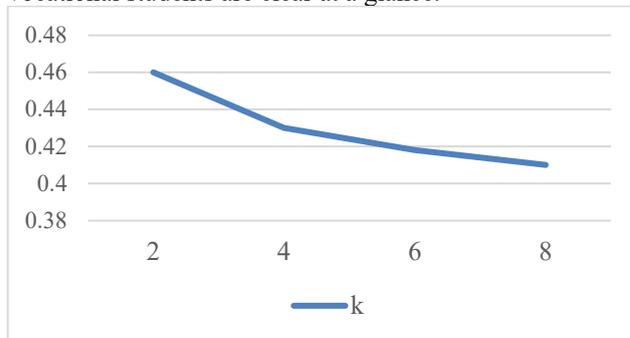


Fig. 1 contour coefficients with different K values

As can be seen from Figure 1, when k = 3, the contour coefficient is the best, but the classification is not detailed enough. In order to make the classification more detailed, take k = 5. It can be seen that the optimized k-means algorithm is relatively stable and trustworthy. The k-means algorithm is used to analyze and process the data. Output to the commonly used Excel tools, the chart can be more easily understand the ideological and political situation of students. The complex algorithm is processed by computer.

Figure 2 shows the experimental analysis of the improved algorithm. Since the data in this paper does not contain target classification, the paper first uses the universal UCI standard data sets iris and wine to evaluate the effect of the improved algorithm. The traditional K-

means algorithm, minmax.k-means algorithm and the method in this paper are used for experiments. The results show that the accuracy of the traditional algorithm and minmax. K-means method is 75.32% and 86.66% respectively, and the contour coefficient is 0.4453 and 0.5049 respectively, while the accuracy of this method can reach 89.32% and the contour coefficient is 0.5356. When using the wine dataset, the accuracy of traditional algorithm and minmax. K-means method is 69.32% and 72.66% respectively, and the contour coefficient is 0.2363 and 0.3049 respectively. When using this method, the accuracy can reach 78.32% and the contour coefficient is 0.3356. It can be seen that the clustering effect of this method is better than the other two algorithms.

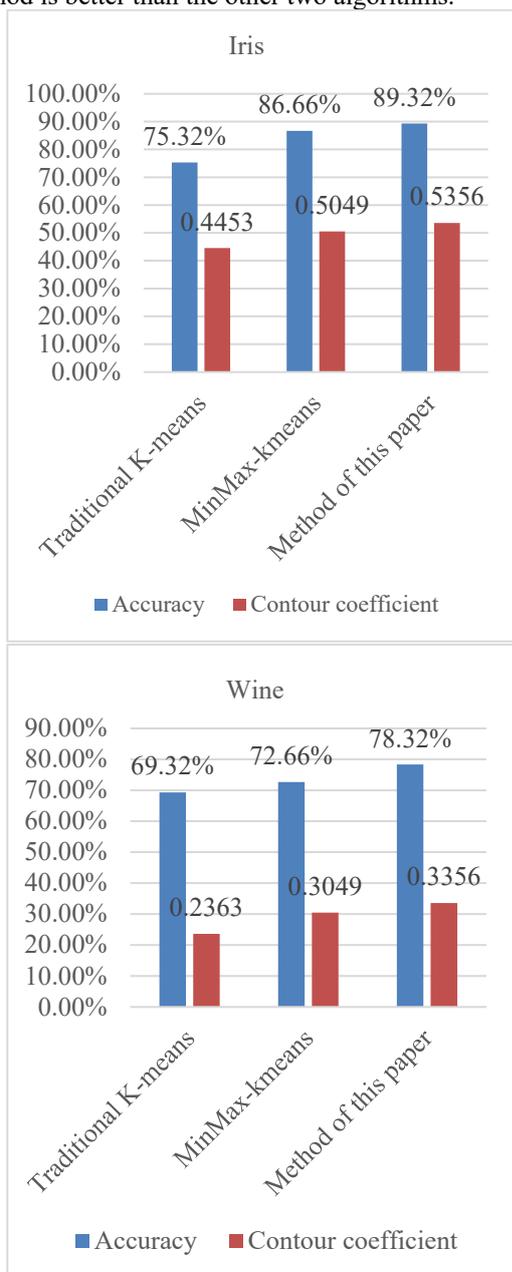


Figure 2 experimental results using UCI data set

5 Conclusion

The assessment function of Ideological and political theory course mainly includes diagnostic function,

guiding function and summarizing function. The teaching goal of Ideological and political theory course determines that the assessment is a comprehensive assessment of students' knowledge, emotion, will and behavior, rather than just one aspect. Therefore, the assessment of Ideological and political theory course should highlight the diagnostic and guiding functions of the assessment, not just reflect the summative function. By giving full play to the diagnostic function of examination, we can help students find and correct the deviation and error in cognition in time, and guide students to establish correct three outlooks. By giving full play to the guiding function of assessment, students' enthusiasm and initiative in learning ideological and political theory course can be stimulated, and ideological and political theory course can become a course that students really like and benefit for life. Therefore, the reform and innovation of the assessment mode of Ideological and political theory course in higher vocational colleges should further clarify and give play to the diagnostic and guiding functions of the assessment of Ideological and political theory course in concept, integrate and make full use of the convenience brought by modern computer technology in technology, and fully develop and utilize the data visualization system, so as to make it the right assistant for teachers and students.

Topic: Scientific research project of Xi'an Aviation Vocational and technical college in 2019: Exploration and practice of model construction of student Party branch in Higher Vocational Colleges in the new era, No.:19XHSK-021.

References

1. Editor. Who's Running the World? Psychological Assessment of Political Leaders[J]. International Bulletin of Political Psychology, 2018, 18(2):2-2.
2. Pasternak A . Political Rioting: A Moral Assessment[J]. Philosophy & Public Affairs, 2018, 46(4):384-418.
3. Sobczyk N . Markey fears 'political interference' on assessment[J]. Greenwire, 2019(MAY 30):12-12.
4. Paul C , Harold K , Chris D , et al. Data Visualization: The Signal and the Noise[J]. IEEE Potentials, 2018, 37(4):28-34.
5. Perkel J M . Data visualization tools drive interactivity and reproducibility in online publishing[J]. Nature, 2018, 554(7690):133-134.
6. Friendly M . HistData: Data Sets from the History of Statistics and Data Visualization[J]. Joint Bone Spine Revue Du Rhumatisme, 2018, 77(3):195-7.
7. Reid, Nathalie. Data Visualization: A Guide to Visual Storytelling for Libraries[J]. Journal of the Medical Library Association: JMLA, 2018, 106(1):135-135.
8. Luo Y , Qin X , Chai C , et al. Steerable Self-driving Data Visualization[J]. IEEE Transactions on Knowledge and Data Engineering, 2020, PP(99):1-1.
9. Amiraghdam A , Diehl A , Pajarola R . LOCALIS:

Locally Adaptive Line Simplification for GPU-based Geographic Vector Data Visualization[J]. *Computer Graphics Forum*, 2020, 39(3):443-453.

10. Avila-Foucat Véronique Sophie, Rodríguez-Robayo Karla Juliana. Determinants of livelihood diversification: The case wildlife tourism in four coastal communities in Oaxaca, Mexico[J]. *Tourism Management*, 2018, 69(DEC.):223-231.