J. COMBIN. MATH. COMBIN. COMPUT. 127b (2025) 2489--2505

Journal of Combinatorial Mathematics
and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Research on the optimization of the ideological education mode in the cultivation of big data talents in colleges and universities based on time series analysis

Juan Zheng[1,✉]

[1] *School of Marxism, Henan Open University, Zhengzhou, Henan, 450046, China*

ABSTRACT
The prediction of the scale of big data talent training in colleges and universities belongs to an important content in the field of big data talent research in colleges and universities. The article uses the primary exponential smoothing method in the time series and the gray model prediction method to predict the scale of college big data talent training and talent demand respectively, and then uses the Lorenz curve and the Gini coefficient to study the matching degree of education in the field of big data. There are experimental results can be obtained, the degree of matching between the professional settings of colleges and universities and the trend of the demand for big data-related positions in enterprises needs to be strengthened, in order to adapt to the future demand for big data-related positions in enterprises, and to further output talents that are in line with the enterprises, the article proposes a model of big data talent cultivation civic and political education in colleges and universities based on the KSAO model. Based on the KSAO model, the ideological education mode of big data talent cultivation in colleges and universities can be implemented at six levels: "theory + project" curriculum system, promoting the dual strategy of "on-campus simulation + off-campus practice", establishing the KSAO multi-dimensional practice assessment system, strengthening the coordination of the industry-teaching cooperation model, building a cloud learning platform with the help of information technology, and implementing the top-down education design.

Keywords: time series, exponential smoothing method, gray model prediction method, civic education mode

1. Introduction

With the rapid development of technology, big data plays a pivotal role in all industries. And in order to effectively utilize big data, the demand for talents is also increasing. In a broad sense, big data talents

✉ Corresponding author.
E-mail address: zjalice625@126.com (J. Zheng).
Received 07 January 2024; Revised 02 August 2024; Accepted 21 November 2024; Published Online 16 April 2025.
DOI: 10.61091/jcmcc127b-141

are scientists and engineers with big data processing ability [1-4]. Currently, a large number of courses in data science, data science degree programs, and data science short-term training courses are offered internationally. From the viewpoint of training programs set up internationally, big data talents should systematically master the skills related to data analysis [5-8]. It mainly includes mathematics, statistics, data analysis, business analysis and natural language processing, etc., and has wider knowledge, strong practical ability, innovation consciousness and teamwork spirit, while the most important point that can not be neglected in the process of big data talent training is the Civic and political education model [9-12].

With the development of the times and social progress, Civic and political education in colleges and universities has become a necessary part of students' comprehensive development. As a kind of practical parenting mode, the ideological and political education in colleges and universities plays an important role in cultivating college students' comprehensive development and success [13-16]. And the ideological and political education model of big data talent cultivation in colleges and universities is a model that shapes students' ideological, moral, and cultural literacy while cultivating their knowledge and ability of big data through various ways, including classroom, campus, and community [17-20]. This model aims to help students grow and exercise in big data teaching practice to achieve their goal of comprehensive growth and all-round development, which includes education methods such as ideological and moral education, education combining big data knowledge and practice, social service and social practice education, etc., and is characterized by diversity, practicality, and relevance [21-24].

Literature [25] emphasized the arduousness of establishing a talent training system for big data technology and suggested that professional development be promoted through solutions such as innovative talent training models based on local needs and improving the structure of the teaching force to have bridging the gap between talents and actual needs. Literature [26] explored the talent training system of data science and big data technology from the aspects of talent training objectives, curriculum system structure, faculty construction, teaching and research reform, etc., aiming to provide reference for the reform and development of talent training for big data majors. Literature [27] aims to build a practical teaching system for financial big data talent cultivation, so as to quickly realize excellent financial big data engineering talents to meet the industry talent demand. Literature [28] emphasizes the wide application of big data technology and the demand for talents in this field, points out that for the cultivation of talents in this field major universities are in the process of continuous practice and exploration, and puts forward relevant strategies to provide reference for the construction of big data majors in colleges and universities. Literature [29] analyzes the talent demand of China's big data industry and the current situation of big data applied talents training in colleges and universities, and puts forward the “three fusion cultivation paths” for big data applied talents in colleges and universities, which includes the “interdisciplinary fusion” and the construction of the talent cultivation environment and other aspects. Literature [30] examined the comprehensive management information system based on the deep integration of industry and education big data talent training model, including the construction of application-oriented innovative talent training system, big data course practice system, government financial support policies. The above studies illustrate the importance of big data talents in the era of extensive application of big data, and the important role played by universities in the cultivation of big data talents, and individual scholars have put forward corresponding strategies by pointing out the current deficiencies in this field, but current academics have not yet carried out a study on the optimization of the Civic and Political Education mode in the cultivation of big data talents.

The article firstly carries out theoretical research on the time series correlation analysis method. Subsequently, the first-order exponential smoothing method is used to analyze the current demand for big data talents, and the observed values are visualized. Then the gray prediction model was used to predict the scale of big data talent training (enrollment scale and student scale) in colleges and universities. By using a large amount of data on China's big data talent and college enrollment scale and graduate supply and demand status was analyzed. Ten colleges and universities were selected for matching research based on the Pearson correlation coefficient and the trending of enterprise job demand. Finally, based on the experimental results, the KSAO model is introduced to propose a model for the cultivation of big data talents in colleges and universities in terms of the curriculum system, strategy and assessment system of the Civic and Political Education.

2. Research methodology

2.1. Fundamentals and Modeling Approach of Exponential Smoothing Models

2.1.1. Time series. Time series refers to the sequence formed by arranging the various values of a certain statistical index of a certain phenomenon at different times in accordance with the time sequence, which is a quantitative forecasting method. And time series analysis is a dynamic statistical method of data processing, based on the theory of stochastic processes and mathematical statistics methods, to study the statistical laws followed by random data series to solve practical problems, which is widely cited in military, economic, weather and other fields. Exponential smoothing is one of the time series algorithms and a common method in production forecasting, and it is also used for short- and medium-term forecasting of economic development trends without considering more distant data, so this study chooses the exponential smoothing method. It is characterized by the fact that S_t is related to all the signals produced before and the closer the signal the more weight it has. It includes first-order exponential smoothing, second-order exponential smoothing, and third-order exponential smoothing, and according to the observation of the research data, it is found that the data do not present a straight line or parabola, so the first-order exponential smoothing method is selected in this study [31]. The basic idea: the predicted value is the weighted sum of previous observations, and different weights are given to different data, with new data given larger weights and old data given smaller weights.

2.1.2. Primary exponential smoothing method. Exponential smoothing is an empirical forecasting method that can be used for short-term forecasting of any kind of time series that do not have an obvious functional pattern, but do have some kind of back-and-forth correlation. The idea comes from an improvement of the moving average method of forecasting. When current and historical values are used to predict future values, the moving average method has two unreasonable situations, one of which is that it is unreasonable to give equal weight to current and historical values. The second is that the window width is difficult to determine, and the observations before the window width have a weight of 0, and the observations within the window width have an equal weight, which is contrary to the actual occurrence of the event. Exponential smoothing, on the other hand, gives each historical value an exponential decay in weight over time, which solves the two unreasonable situations of the moving average method. Exponential smoothing method is divided into various types, commonly used simple exponential smoothing method, double exponential smoothing method.

Exponential smoothing method is proposed by Brown, he believes that the time series of the trend has stability or regularity, so the time series can be reasonable to follow the trend of deferral. He believed that the recent past dynamics, to some extent, will continue into the recent future, so larger weights are placed on the most recent information. The fundamental purpose of the exponential smoothing method is to remove some of the random fluctuations so as to find the obvious regularities in them and make reasonable predictions of future trends. The method is based on the sliding average method and is divided into primary smoothing, secondary smoothing, and tertiary smoothing.

Let $x_1, x_2, \cdots, x_n$ be an observation $(t = 1, 2, \ldots, n)$ at time t that is independent and follows a normal distribution $N(0, \sigma^2)$ and, for the general case, do substitutions $y_t = x_t - \mu$ and $S_t^{(1)}$ for the definition of a time series in which time t reaches one exponential smoothing value:

$$S_t^{(1)} = \alpha x_t + (1-\alpha) S_{t-1}^{(1)} (t = 1, 2, \cdots, n, 0 \le \alpha < 1) \tag{1}$$

where $S_t^{(1)}$ is the smoothed value of time t. x_t is the actual value of time t. $S_{t-1}^{(1)}$ is the smoothed value of time $t-1$. α is the smoothing constant, which takes values in the range [0, 1]. $S_t^{(1)}$ is the weighted arithmetic mean of x_τ and $S_{t-1}^{(1)}$, with α taking the size of the change in the size of the decision x_z and $S_{t-1}^{(1)}$ on the degree of influence of $S_t^{(1)}$, when a takes 1, $S_\tau^{(1)} = x_z$. $s_t^{(1)} = S_{t-1}^{(1)}$ when α takes 0.

This is obtained from the recurrence relation:

$$\begin{aligned} S_t^{(1)} &= \alpha x_t + (1-\alpha)\left[\alpha x_{t-1} + (1-\alpha) S_{t-2}^{(1)}\right] \\ &= \alpha x_t + \alpha(1-\alpha) x_{t-1} + (1-\alpha)^2 \left[\alpha x_{t-2} \div (1-\alpha) S_{t-3}^{(1)}\right] \\ &= \alpha x_t + \alpha(1-\alpha) x_{t-1} + (1-\alpha)^2 \left[\alpha x_{t-2} + \cdots + \alpha(1-\alpha)^{t-1} x_1 + (1-\alpha)^t S_0^{(1)}\right] \end{aligned} \tag{2}$$

Since $0 \le \alpha < 1$, so $t \to \infty$ when $\lim_{t\to\infty} (1-\alpha)^t = 0$

Then:

$$S_t^{(1)} = \alpha x_t + \alpha(1-\alpha) x_{t-1} + \alpha(1-\alpha)^2 x_{t-2} + \cdots + \alpha(1-\alpha)^{t-1} x_1 \tag{3}$$

It can be said that $S_t^{(1)}$ is the exponentially weighted average of the observations of period t and previous periods, and the weights of the observations decrease geometrically according to the recursive period, and the further away the data of each period is from period t, the smaller its coefficient is, and thus the smaller is its impact on the predicted value.

One-time exponential smoothing method, also known as simple exponential smoothing method, which is applicable to smooth series without linear trend or seasonal variation. Simple exponential smoothing method for forecasting is actually based on the current forecast error of the current forecast value to make certain adjustments to get the next forecast value, that is:

The new forecast = the old forecast + α × the old forecast error, the magnitude of the adjustment to the old forecast depends on the size of the α, the larger the α, the greater the adjustment made, the smaller the α, the smaller the adjustment made.

2.2. *Gray prediction model*

The research object of grey system theory is the "small sample" and "poor information" uncertain system with "partial information known, partial information unknown", which mainly generates and develops "partial" known information, extracts the required information, and completes the relevant prediction and effective control of the time-related gray process that changes within a certain range of the system. The basic idea is to reduce the gray scale of the system by accumulating new sequences of the observed data series according to certain rules, so as to discover the latent order rules of the system and construct a model [32]. The commonly used models for grey systems are univariate first-order models and multivariate first-order models. The grey system model used in this paper is a univariate first-order model, i.e., the GM(1,1) model. Figure 1 shows the meaning of the model symbol.

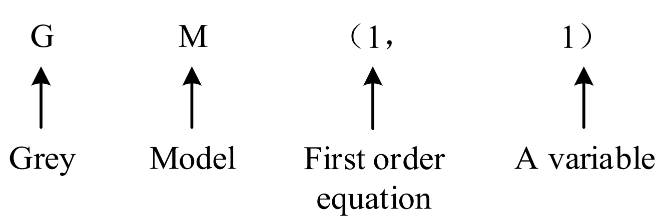


Fig. 1. Schematic diagram of GM (1,1) in gray model

2.2.1. Basic Ideas:

1) Derive the original form of the gray model GM(1, 1). Let $X^{(0)}$ be the original data column and $X^{(0)}=(x^{(0)}(1),x^{(0)}(2),\ldots,x^{(0)}(n))$, where $x^{(0)}(k)\geq 0,k=1,2,\cdots,n$. Is cumulated over the original data column:

$$\begin{gathered} x^{(1)}(1)=x^{(0)}(1) \\ x^{(1)}(2)=x^{(0)}(1)+x^{(0)}(2) \\ \cdots \\ x^{(1)}(k)=\sum_{i=1}^{k}x^{(0)}(i),k=1,2,\cdots,n \end{gathered} \tag{4}$$

This yields the new data sequence $X^{(0)}$, i.e., $X^{(1)}=(x^{(1)}(1),x^{(1)}(2),\cdots,x^{(1)}(n))$, $X^{(1)}$ is called the 1-AGO (one-aggregate generating operator) sequence for $X^{(0)}$.

The original form of the GM(1, 1) model for the gray system is derived from the original series and the 1-AGO (one-accumulation generating operator) sequence:

$$x^{(0)}(k)+ax^{(1)}(k)=b \tag{5}$$

Where a,b is the gray system parameters to be determined, a is the development coefficient, reflecting the development of the series, and b is the gray system action quantity, reflecting the data change relationship [33].

2) Determine the basic form of the gray model GM (1, 1) Let Z (1) be the immediate neighboring mean generating sequence of X (1), that is:

$$Z^{(1)}=(z^{(1)}(2),z^{(1)}(3),\cdots,z^{(1)}(n)) \tag{6}$$

Among them,

$$z^{(1)}(k)=0.5((x^{(1)}(k)+x^{(1)}(k-1)),k=2,3,\cdots,n \tag{7}$$

The basic form of the gray model GM(1, 1) is derived:

$$x^{(0)}(k)+az^{(1)}(k)=b \tag{8}$$

2.2.2. **Solving for gray system parameters a and b.** If $\hat{a}=[a,b]^T$ is a parameter column, then the least squares estimation of the parameter columns of the GM(1, 1) model $x^{(0)}(k)+az^{(1)}(k)=b$ is satisfied:

$$\hat{a}=(B^TB)^{-1}B^TY \tag{9}$$

Matrix Y and Matrix B respectively:

$$Y=\begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}, B=\begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix} \tag{10}$$

Substituting matrix Y and matrix B into Eqs. To find a and b.

2.2.3. Deriving model equations. Confirm that the whitening equation of the gray differential equation $x^{(0)}(k)+az^{(1)}(k)=b$ is also known as the shadow equation.

$$\frac{dx^{(1)}}{dt}+ax^{(1)}=b \tag{11}$$

The solution of the whitening equation is solved to obtain the time response sequence:

$$\hat{x}^{(1)}(k+1)=(x^{(0)}(1)-\frac{b}{a})e^{-ak}+\frac{b}{a};k=1,2,\cdots,n \tag{12}$$

And then by:

$$\hat{x}^{(0)}(k+1)=\hat{x}^{(1)}(k+1)-\hat{x}^{(1)}(k) \tag{13}$$

Cumulative subtraction reduction is obtained:

$$x^{(0)}(k+1)=a^{(0)}\hat{x}^{(1)}(k+1)=(1-e^{a})(x^{(0)}(1)-\frac{b}{a})e^{-ak};k=1,2,\cdots,n \tag{14}$$

2.2.4. Derive and test the model data. Use the formula to find out the data to be simulated, use the residuals and the relative value of the residuals to test the accuracy of the simulated data, the smaller the relative value of the residuals indicates that the model is more accurate.

1) Residual series:

$$\begin{aligned}\varepsilon^{(0)}&=(\varepsilon^{(0)}(1),\varepsilon^{(0)}(2),\ldots\varepsilon^{(0)}(n))\\&=(x^{(0)}(1)-\hat{x}^{(0)}(1),x^{(0)}(2)-\hat{x}^{(0)}(2),\ldots,x^{(0)}(n)-\hat{x}^{(0)}(n))\end{aligned} \tag{15}$$

2) Sequence of residual relative values:

$$\Delta=(\frac{\varepsilon(1)}{x^{(0)}(1)},\left|\frac{\varepsilon(2)}{x^{(0)}(2)}\right|,\cdots,\left|\frac{\varepsilon(n)}{x^{(0)}(n)}\right|)=\{\Delta_k\}_1^n \tag{16}$$

For $k\le n$, call $\Delta_k=\left|\frac{\varepsilon(k)}{x^{(0)}(k)}\right|$ the k-point simulated relative error and call $\bar{\Delta}=\frac{1}{n}\sum_{k=1}^{n}\Delta_k$ the average relative error. A residual pass model is, given α, when $\bar{\Delta}<\alpha$ and $\Delta_n<\alpha$. Given the value of α, one level of test model simulation accuracy can be determined. There are four commonly used modeling accuracy classes as follows: when $\alpha=0.01$, the model accuracy is Class I. When $\alpha=0.05$, the model accuracy is Level II. When $\alpha=0.10$, the model accuracy is class III. When $\alpha=0.20$, the model accuracy is class IV.

2.3. Lorenz curve and Gini coefficient

The Lorenz curve responds to the degree of imbalance in the distribution of income in a country or region [34]. The Lorenz curve is shown in Figure 2. The horizontal axis of the graph represents the cumulative percentage of population and the vertical axis represents the cumulative percentage of income.OI represents the absolute average of income and is called the absolute mean line.Any point on this line indicates that a given percentage of the population, receives the same percentage of income in the distribution of income.OXI is known as the absolute unevenness line and indicates that of all the members all but one have zero income.OFGI represents the unevenness line of the income, i.e., the actual income distribution curve, the larger the difference from the OI line, the more uneven it is.

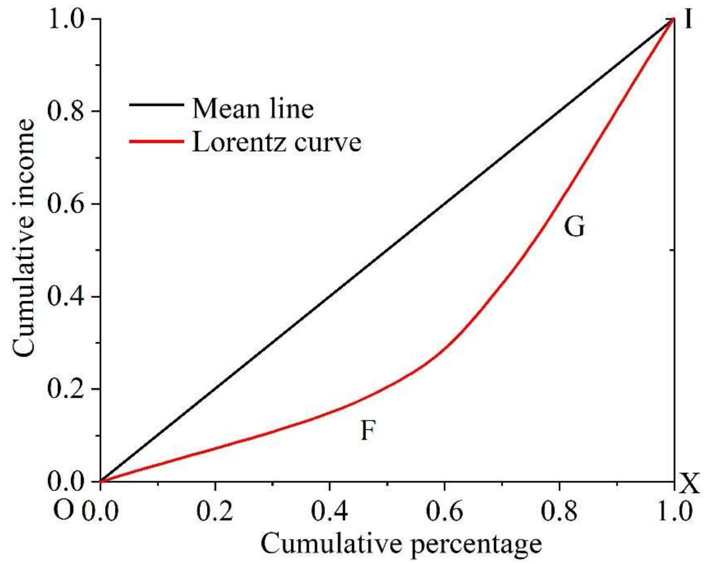


Fig. 2. Lorentz curve

The Italian economist Gini further quantified the Lorenz curve by proposing the Gini coefficient, also known as the Lorenz coefficient, which is generally denoted by G:

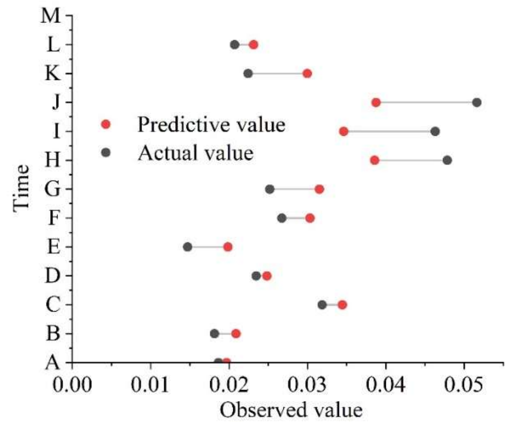
$$G = A/(A+B) \tag{17}$$

where: A denotes the area between the Lorenz curve and the absolute average. B represents the distance between the Lorenz curve and the absolute non-average. The value of $A+B=1/2$, G is [0,1], the larger the G value, the more uneven the income distribution, and vice versa, generally according to international standards and practices, the Gini coefficient is less than 0.2 for the "absolute average", 0.2~0.3 for the "relative average", 0.3~0.4 for the "relatively reasonable", 0.4~0.5 for the "gap is too large", and above 0.5 for the "high uneven", so 0.5 is the warning line.

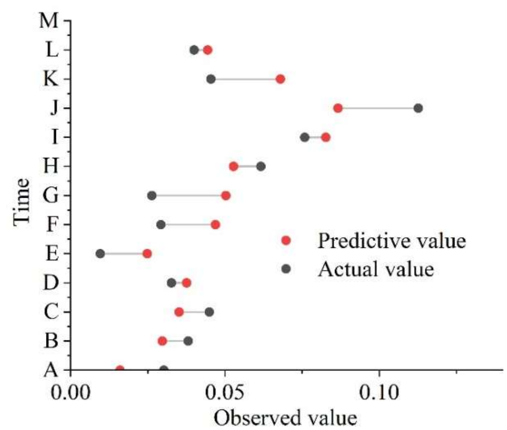
In summary, the Lorenz curve and the Gini coefficient are objective indicators used to measure the degree of imbalance in the distribution of income, but they are essentially studies of the imbalance in the distribution of resources. Therefore, using the Lorenz curve and the Gini coefficient to describe the degree of non-equilibrium in the distribution of educational resources is in good agreement. This paper draws on the principles of the Lorenz curve and the Gini coefficient to construct an educational resource matching model, seeking a new quantitative indicator for the degree of imbalance in the distribution of educational resources in the field of big data.

3. Analysis of big data talent needs

In this section, we select the seven keywords with the highest relevance to big data job categories: SLAM, SPARK, Kaldi, CUDA, Storm, NLP, and OpenGL, and study the demand trend of these seven highly relevant skill keywords in each quarter from 2021 to 2023 according to the first-order exponential smoothing method, and predict the demand trend in the next quarter. In order to better interpret the prediction results, the results are visualized in this section. Figure 3 shows the trend of OPENGL and SLAM skill prediction (Figure a is the trend chart of OPENGL skill prediction, and Figure b is the trend chart of SLAM skill prediction). A~M represents Q1 2020, Q2 2020, Q3 2020, Q4 2020, Q1 2021, Q2 2021, Q3 2021, Q4 2021, Q1 2022, Q2 2022, Q3 2022, Q4 2022, and Q1 2023, respectively. SLAM and OPENGL skills are high relevant skills for the training of big data talents in colleges and universities respectively, as can be seen from the figure, the demand for these two skills fluctuates a lot in three years, and in the last half of the year, the demand began to gradually decrease, and for the forecast value of the first quarter of 2023 is also a decrease in the percentage, but the overall relative frequency has increased, and as can be seen from the previous section, the demand for these two types of positions in the last three years overall rise and The proportion is relatively high, so the proportion of SLAM and OPENGL skills will also increase with the increase in demand for job categories.



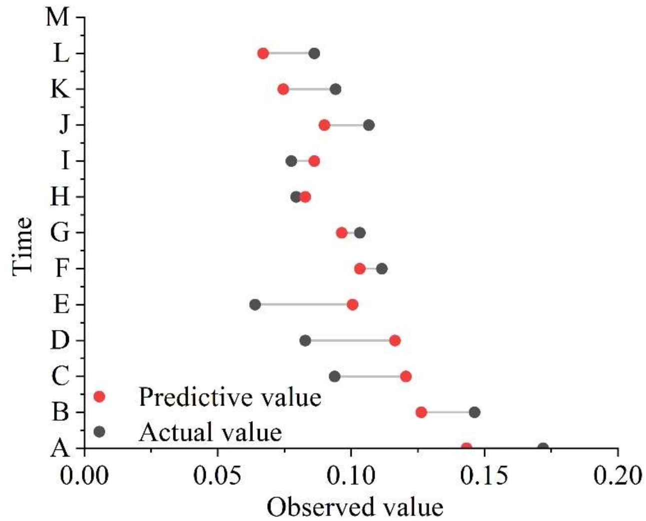
(a) OPENGL



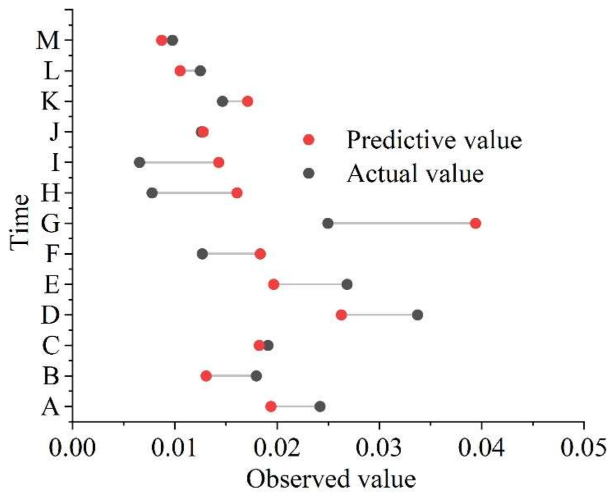
(b) SLAM

Fig. 3. OPENGL and SLAM skills forecasting trends

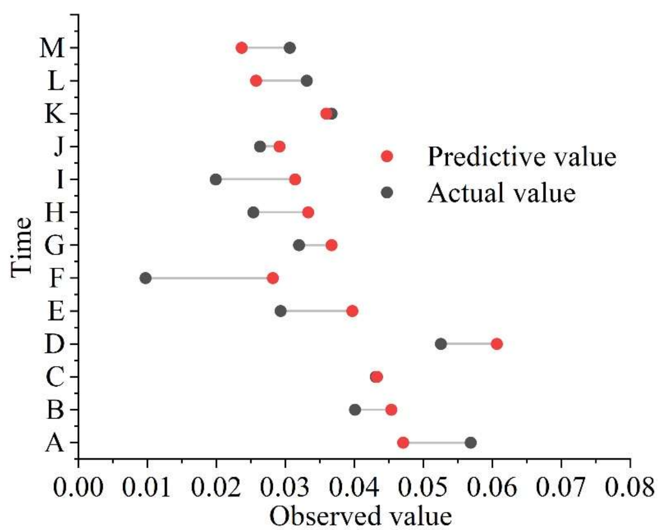
SPARK, KALDI, CUDA and STORM skills need forecast trends are shown in Figure 4 (Figures a~d show SPARK, KALDI, CUDA and STORM skills need forecast trends, respectively). The trend line of the figure shows that the demand for SPARK, KALDI, CUDA and STORM shows a straight line downward trend, and the forecast demand in the next quarter will also be slightly reduced in the proportion of the SPARK skills it belongs to the machine learning category of the core skills, is a can be used for large-scale data processing of a common engine (core). Because it is based on memory computing, so once the amount of data beyond the memory is unstable, and its own task scheduling function is poor, with the machine learning algorithms of the iterative upgrading, there will be more stable data processor to replace it. The demand for CUDA and STORM and KALDI skills with the quarter time and reduce, not only because of the job category to which it belongs to the demand also has a slow decline, the software's own limitations will also affect the demand for its share of the reduction for job seekers, the core skills must be required to master, but in the discovery of skills trends and technology updates at the same time, it is also necessary to identify the emergence of new technologies. However, while identifying trends in skill changes and technological upgrades, it is also important to recognize the emergence of new technologies so as to master and learn new skills.



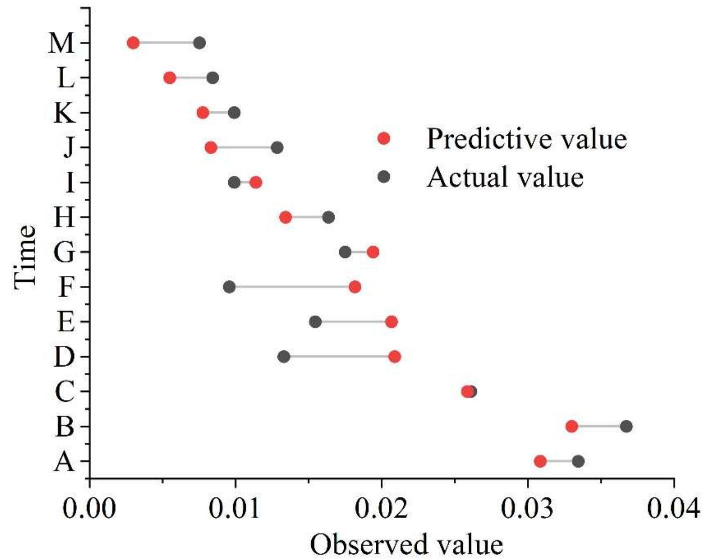
(a) SPARK



(b) KALDI



(c) CUDA



(d) STORM

Fig. 4. Predict Trends

The NLP skills need forecast trend is shown in Figure 5. It can be seen from the figure that the predicted trend line for NLP skills shows large fluctuations.The reason for the high and low demand for NLP can be learned from the decrease or increase in hiring for jobs in the Natural Language Processing category, which has the second highest demand in 2020, with a share of 15%, and then decreases by almost half in 2021 and 2022, with a consequent decrease in hiring demand.

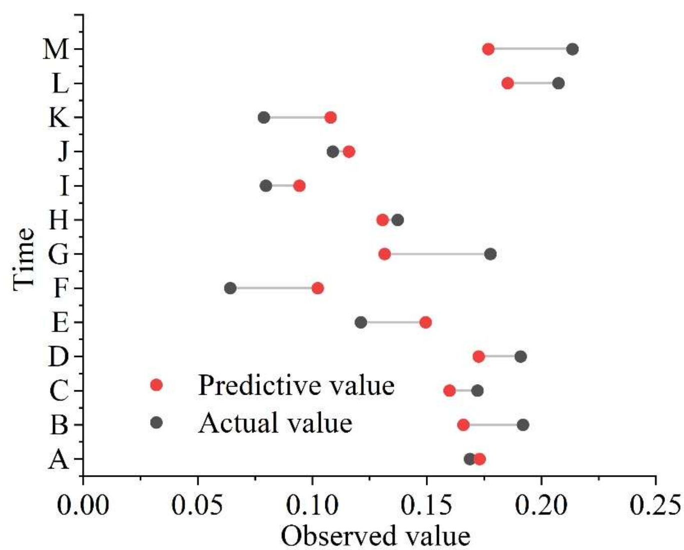


Fig. 5. NLP skills need to predict trends

4. Forecast of the scale of big data talent training in higher education institutions

This section uses the gray prediction model to synthesize the historical data with the prediction results, and the development trend of big data talent cultivation scale in colleges and universities from 2016 to 2023 can be mapped.The historical data of big data talent scale in colleges and universities is shown in Table 1.The development trend of big data talent cultivation scale in colleges and universities from 2010 to 2023 is shown in Fig. 6.From 2016 to 2023, the development trend of big data talent cultivation scale in China's The scale of big data talent cultivation in colleges and universities increases year by year, and the average annual growth rate of enrollment scale is 16%, and the average annual growth rate of enrolled students scale is 16.6%.

Table 1. The number of universities is historical data

Statistical category	Talent scale history data							
	2016	2017	2018	2019	2020	2021	2022	2023
Enrollment scale/person	12095	14820	15754	17967	25518	28455	31321	33145
Annual growth rate/%	—	22.5	6.3	14.0	42.0	11.5	10.1	5.8
Annual growth rate/%	16.0							
Student size/person			49565	56232	69575	82612	97303	106291
Annual growth rate/%	—	—	—	13.5	23.7	18.7	17.8	9.2
Annual growth rate/%	16.6							

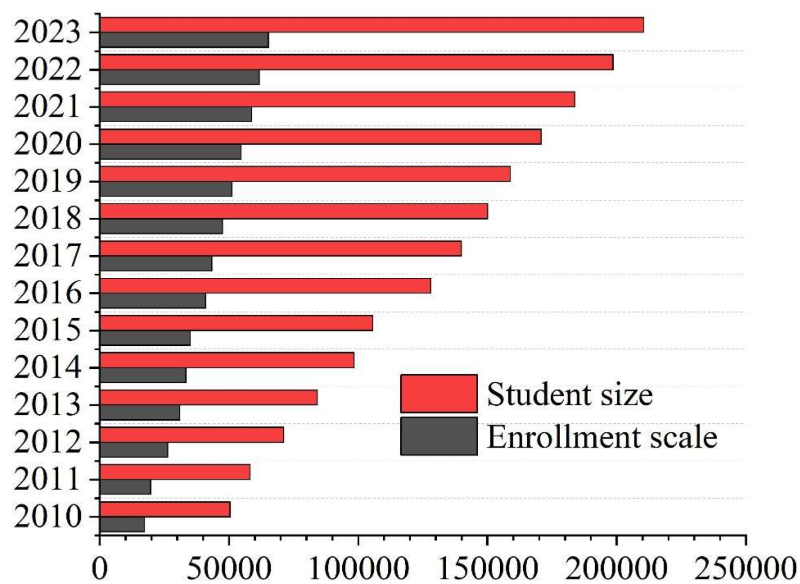


Fig. 6. The number of universities in 2010-2023 is based on the trend of talent scale

5. Study on the match between supply and demand of big data talents

This section borrows the Lorenz curve graph for research and calculates the Gini coefficient on its basis to quantitatively describe the matching situation between the status of new big data majors and the education level of local colleges and universities in each province of the country.

Using x to represent the cumulative percentage of the number of regional college construction, and y to represent the percentage of the corresponding big data major new additions, the Lorenz curve graph of regional college major setting and education level is plotted, and the Lorenz curve of regional college major setting and education level is shown in Figure 7.

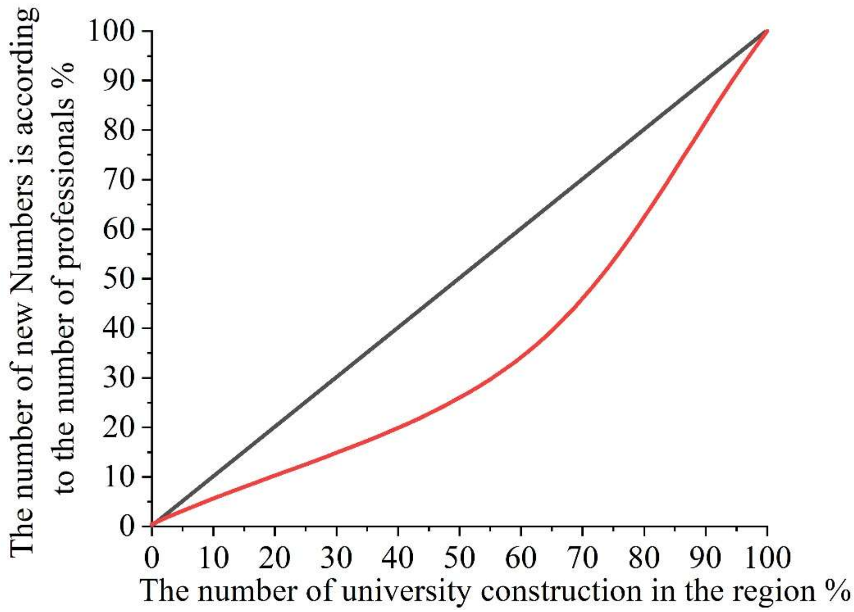


Fig. 7. Regional university professional Settings and education level lorentz curve

The change in the number of college majors set up is determined by the Ministry of Education and national related policies comprehensively every year, while the demand for enterprise talents comes from the market demand, so the variables are independent of each other, in line with the model constraints. Since the correlation coefficient method is only sensitive to the correlation of observations, and is not sensitive to the size of observations, the number of people who set up the specialty in ten colleges and universities in 2016-2023 is Y, Y=[8122,9642,9233,9154,9563,9647,9796,10055]. Considering that colleges and universities need to train talents for a period of four years, it is agreed that the number of colleges and universities specializing in setting up the number of people corresponds to the amount of enterprise demand in 2016 for the model assessment, and the correspondence of variables in other years and so on, so that the amount of colleges and universities specializing in setting up the number of people corresponds to the amount of enterprise demand for the predicted number of people recruited in the years 2016-2023, and using the method of this paper to get the talent demand of the enterprises related to big data for the next four years predicted value, define the peak of enterprise demand for big data-related talent each year as the enterprise talent demand in that year, noted as X, then X = [205,229,273,372,484,585,752,936]. Using SPSS software to draw a scatter plot of the amount of college specialization and enterprise demand, the amount of college specialization and enterprise demand is shown in Figure 8. From the figure, it can be obtained that the scatter plot of the amount of college and university big data major settings and enterprise demand tends to a straight line, and the sequence has obvious correlation. Now the relationship is further quantified by Pearson correlation coefficient method. The analysis results show that the observed values X and Y have positive correlation, and the degree of correlation between the two is medium, that is, the degree of matching between the college professional settings and the trend of enterprise big data-related job demand needs to be strengthened, in order to adapt to the future demand of big data-related jobs in enterprises, and to further output talents that are in line with the enterprises, the professional training program can be optimized appropriately.

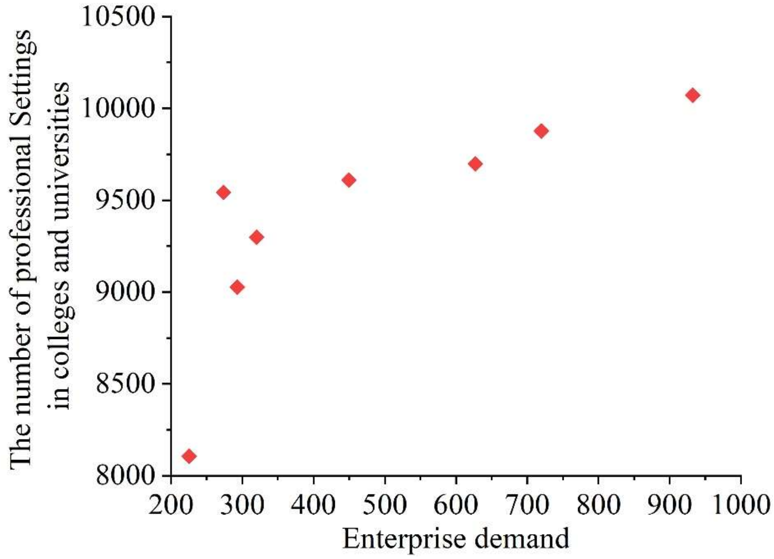


Fig. 8. The quantity of professional Settings and the demand of enterprises

6. Civic and political education model of big data talent cultivation in colleges and universities based on the KSAO model

The introduction of the KSAO model allows for a more comprehensive analysis of the vocational competencies necessary for an individual to engage in a certain occupation. According to the model to analyze the abilities required for the relevant positions in which big data students are engaged, the following is a specific analysis of the vocational abilities required for big data professionals. In view of the results of the current research on the cultivation of big data talents in colleges and universities, the KSAO model is proposed as a model for the cultivation of big data talents in colleges and universities based on the KSAO model of the ideological and political education, the construction of a scientific discipline system, the enhancement of training in solving practical problems, and the focus on the cultivation of vocational development and lifelong learning ability.

6.1. Building a "theory + project" curriculum system

According to the international trade project process, the theoretical learning in class lays the foundation for the basic vocational practice, and the theoretical learning results are tested through actual projects. Specifically manifested in the setting of "enterprise product promotion training" and "foreign trade documents confirmed training" and other on-campus practical content, off-campus business transactions real case analysis. The purpose is to familiarize students with the complete foreign trade process, master the relevant financial management, customer management and the production of documents and other theoretical foundations, in accordance with the current industry standards to strengthen the understanding of international trade documents, and quickly improve the ability to integrate into enterprises. In addition to basic theoretical knowledge, diversified computer skills classes are offered.

In the daily theoretical learning, the theoretical knowledge is used as a carrier, and the elements of ideology and politics are explored to cultivate dialectical thinking. Understand the social responsibility of enterprises in international trade. Guide students to pay attention to the impact of different cultures on international trade and how to realize win-win cooperation in cultural differences, and integrate implicit ideological education into explicit theoretical knowledge. In the classroom, we adopt the chain learning method of teacher's lecture, class discussion and post-class reflection to strengthen students' self-reflection and life value reshaping in addition to theoretical knowledge, and to realize the deeper meaning of their majors.

6.2. *Promote the dual strategy of "on-campus simulation + off-campus practice"*

1) On-campus simulation. Enriching practical teaching methods, simulation training teaching is carried out by teacher-student interaction and peer interaction. Through the use of simulation software for transaction deduction, big data students can experience the processes of communication with customers, negotiation with suppliers, financial transactions with banks and coordination with freight forwarding intermediaries in the on-campus simulation environment, and at the same time, expand their perspective on commodity trading concerns.
2) Off-campus practice. Through off-campus practicums, students are able to directly face market competition and business challenges, and interact in an actual trading environment. Meanwhile, exchanges with industry professionals help students understand the latest trends and practical experience in the industry and apply them to actual projects.

6.3. *Establishment of KSAO multidimensional practice assessment system*

The course assessment is based on theoretical knowledge as the main assessment content, in addition to examining class attendance and in-class interaction. In addition, students' professional value identity, professional ethics and case values related to professional projects are analyzed to provide feedback on the strengths and weaknesses of the course Civics in professional knowledge learning. The internship assessment is jointly conducted by the instructor and the enterprise supervisor, and is based on multi-dimensional considerations such as internship attitude, professionalism, internship report and internship results. The evaluation of practical ability is categorized into four fuzzy evaluation levels, namely "excellent", "good", "average" and "poor". This helps students to better recognize the level of their practical ability.

6.4. *Strengthening the harmonization of models of cooperation between industry and education*

As an important force to promote the development of local economy, local colleges and universities should strengthen the cooperation and exchange with enterprises. Timely collection of industry development dynamics, combined with the current employment environment to develop student training programs. Firstly, invite excellent enterprises outside the school to hold lectures in the school to introduce the prospects of the industry, as well as to conduct training to cultivate vocational concepts before entering the workforce. Secondly, we jointly build training bases with schools and enterprises and establish industrial incubators to provide teachers and students with a practical platform that is closer to the real trading environment, and at the same time, help solve some of the students' internship problems and enrich their practical experience. Finally, collect data on the current foreign trade industry environment. For the training of applied talents, the smooth integration into the job market is the feedback of whether the talent training program is practical standard, improve the job market data collection and analysis, is an important means of practical education and training. In addition to traditional practice and simulation training, students are motivated to participate in professional competitions such as innovation and entrepreneurship, and students are organized to challenge foreign trade skill competitions with significant influence within the big data profession as well as scientific research projects, in order to further promote the improvement of students' practical ability.

6.5. *Information Technology Enabled Cloud Learning Platforms*

The use of modern digital information technology promotes the diversification of education modes and realizes the opening and sharing of educational resources. The cloud learning platform is based on multimedia development in the Internet era, breaking the time and space limitations of traditional classroom teaching, enriching students' access to information, and providing a new position for practical education. Utilizing the network platform, constructing new media for the circulation of information on international trade, building WeChat public number and short videos for knowledge sharing, forming a dynamic popularization of information on international trade and a knowledge platform for teachers and students to participate in together. Popularize the awareness of responsibility and network security in the practice process of international trade, cultivate students' national security concepts, and shape road confidence, theoretical confidence, institutional confidence and cultural confidence. Help students resist bad temptations and enhance the extensibility of the course's civic politics.

7. Conclusion

This study applies the time series analysis method to the prediction of big data talent cultivation in colleges and universities, and explores the time series of the prediction of big data talent cultivation in colleges and universities in the time series methodology, and provides theoretical basis for the development of scientific and effective civic education model from the theoretical point of view. This paper draws the following conclusions:

1) In a trend analysis of Big Data talent supply and demand forecasts for 2016-2023. During the period of 2016-2023, the difference between the demand for big data talents and the supply E>0, the total supply is smaller than the total demand, and the gap between supply and demand by 2023 is about 10,000 people.

2) The scale of big data talent training in colleges and universities increased year by year from 2016 to 2023, and the average annual growth rate of enrollment size was 16%, and the average annual growth rate of enrolled student size was 16.6%.

About the Author

Juan Zheng was born in Xuchang, Henan, P.R. China, in 1981. She obtained a master's degree from South China University of Technology in China. I am currently working at the School of Marxism, Henan Open University. My main research direction is Ideological and political education and Higher education and management.

References

[1] Oussous, A., Benjelloun, F. Z., Lahcen, A. A., & Belfkih, S. (2018). Big Data technologies: A survey. Journal of King Saud University-Computer and Information Sciences, 30(4), 431-448.

[2] Hashem, I. A. T., Chang, V., Anuar, N. B., Adewole, K., Yaqoob, I., Gani, A., ... & Chiroma, H. (2016). The role of big data in smart city. International Journal of information management, 36(5), 748-758.

[3] Dash, S., Shakyawar, S. K., Sharma, M., & Kaushik, S. (2019). Big data in healthcare: management, analysis and future prospects. Journal of big data, 6(1), 1-25.

[4] Qaffas, A. A., Ilmudeen, A., Almazmomi, N. K., & Alharbi, I. M. (2023). The impact of big data analytics talent capability on business intelligence infrastructure to achieve firm performance. foresight, 25(3), 448-464.

[5] Jin, Y., Wang, Q., Wang, Y., Geng, K., Ren, Y., & Wei, Y. (2020). Talent Training Mode of Big Data Marketing Major in Higher Vocational Colleges Based on: Scientific Research Assistance, Competition Driven. Education and Training Combination, 3(3).

[6] Liu, Y. (2021). Research on Practical Teaching System of Data Science and Big Data Technology Specialty Based on Application-Oriented Talent Training. Learning & Education, 10(7), 237-238.

[7] Na, Y., & Yuxiang, H. (2020, March). Thoughts on Talent Training of Data Science and Big Data Technology in Local Agricultural Colleges—Take Tianjin Agricultural University as an example. In International Conference on Modern Educational Technology and Innovation and Entrepreneurship (ICMETIE 2020)

(pp. 476-479). Atlantis Press.

[8] Li, H., & Li, Y. (2023). Exploring the Cultivation of Innovative Talents in the Era of Big Data and Cloud Computing. Adult and Higher Education, 5(1), 19-27.

[9] Huang, J. C., & Ko, K. M. (2020). Cultivation of Innovative talents in higher education based on big data. In Data Processing Techniques and Applications for Cyber-Physical Systems (DPTA 2019) (pp. 177-182). Springer Singapore.

[10] Akter, S., Fosso Wamba, S., Barrett, M., & Biswas, K. (2019). How talent capability can shape service analytics capability in the big data environment?. Journal of Strategic Marketing, 27(6), 521-539.

[11] Russell, C., & Bennett, N. (2015). Big data and talent management: Using hard data to make the soft stuff easy. Business Horizons, 58(3), 237-242.

[12] Yang, J., Lyu, X., Xu, X., Jiang, M., & Lyu, P. (2023). Does big data application improve Chinese firms' innovation: the role of talent accessibility. Applied Economics Letters, 1-5.

[13] Damiani, V., & Fraillon, J. (2025). Civic and citizenship education, global citizenship education, and education for sustainable development: an analysis of their integrated conceptualization and measurement in the international civic and citizenship education study (ICCS) 2016 and 2022. Large-scale Assessments in Education, 13(1), 3.

[14] Wang, J., & Zhu, M. (2022). Research on the construction of college English reading courses based on ideology and politics curriculum. Scientific and Social Research, 4(1), 181-185.

[15] Wang, B., Yu, H., Sun, Y., Zhang, Z., & Qin, X. (2023). An effect assessment system for curriculum ideology and politics based on students' achievements in Chinese engineering education. International Journal of Advanced Computer Science and Applications, 14(1).

[16] Wang, T., Fan, D., & Dai, F. (2021). The practice of integrating curriculum ideology and politics into university physics course teaching organically. International Journal of Higher Education Teaching Theory, 2(6).

[17] Meng, F. (2021). Research on the teaching reform of college English translation under the background of curriculum ideology and politics. Journal of Contemporary Educational Research, 5(12), 76-81.

[18] Wang, C. (2023). The Educational Path of Curriculum Ideology and Politics in Chinese International Education. Curriculum and Teaching Methodology, 6(2), 59-65.

[19] Santoso, G. (2019). Philosophical curriculum of civic education in 1975-2013 in indonesia. SEMNASFIP.

[20] Huang, S., Fang, H., & Zhang, Y. (2023, June). Reflection on the Current Curriculum Ideology and Politics in Private Undergraduate Colleges in Guangdong Province from the Perspective of Students. In 2023 4th International Conference on Education, Knowledge and Information Management (ICEKIM 2023) (pp. 497-509). Atlantis Press.

[21] Xie, H., & Song, J. (2022). RESEARCH ON THE GUIDING ROLE OF CURRICULUM IDEOLOGY AND POLITICS ON STUDENTS'SOCIAL CHARACTER. Psychiatria Danubina, 34(suppl 4), 818-818.

[22] Danrong, M. (2023). A study on strategies for improving the professional quality of public English teachers under the background of curriculum ideology and politics. Frontiers in Educational Research, 6(21).

[23] Wu, J., & Wang, X. (2022). Research on the Mode of Specialized and Integrated Education in Colleges and Universities Based on the Perspective of Curriculum Ideology and Politics. International Journal of Education and Humanities, 3(1), 66-69.

[24] Gong, Y., Lu, X., & Cui, D. (2021). Research on the Construction Strategies of" Curriculum Ideology and

Politics" for Professional Courses. International Journal of Frontiers in Sociology, 3(8.0).

[25] Wanga, R., Zhangb, Y., Gec, M., Chend, C., Hane, L., Liuf, H., ... & Zhangh, K. (2022). Big-Data-Technology Talent Cultivation in the Application-oriented University: Practice and Prospects with a Case Study. International Journal of Frontiers in Sociology, 4(5).

[26] Li, M., Yueli, D., Yunfei, Z., & Yingdong, X. (2021, August). Research on professional talent training mode on data science and big data technology in local application-oriented universities. In 2021 International Conference on Big Data Engineering and Education (BDEE) (pp. 56-59). IEEE.

[27] Maalla, A., & Lin, S. Q. (2018, November). The practice teaching system of big data talent training in universities. In 3rd Annual International Conference on Education and Development (ICED 2018) (pp. 159-165). Atlantis Press.

[28] Jiajia, D. (2020). Research on Talent Training Mode of Applied Big Data Specialty in Higher Vocational Colleges. Frontiers in Educational Research, 3(13), 150-155.

[29] Ju, X. (2022). Research on the Training Path of Big Data Application-Oriented Talents in Chinese Colleges and Universities. In Advances in Smart Vehicular Technology, Transportation, Communication and Applications: Proceedings of VTCA 2021 (pp. 411-420). Springer Singapore.

[30] Cuiling, Z., & Chengwei, Y. (2021, November). Explore the mode of big data talent training based on integration between industry and education for IMIS. In 2021 11th International Conference on Information Technology in Medicine and Education (ITME) (pp. 586-590). IEEE.

[31] Herman Lule,Micheal Mugerwa,Robinson Ssebuufu,Patrick Kyamanywa,Jussi P. Posti & Michael Lowery Wilson. (2025). Prospective multi-centre analysis of rural trauma team development training for medical trainees and traffic law enforcement professionals in Uganda: an interrupted time series study. BMC Medical Education(1),148-148.

[32] Liu Caixia,Xu Zhenguo,Zhao Keyun & Xie Wanli. (2023). Forecasting education expenditure with a generalized conformable fractional-order nonlinear grey system model.. Heliyon(6),e16499-e16499.

[33] Hui-Wen Vivian Tang & Tzu-chin Rojoice Chou. (2016). On the fit and forecasting performance of grey prediction models for projecting educational attainment. Kybernetes(9),1387-1405.

[34] Xavier Harmony. (2024). Docked Bikeshare Equity and Goal Conflict: An Evaluation Using Gini Coefficients and Lorenz Curves. Public Works Management & Policy(4),561-587.