

Data Analysis Method of Innovation and Entrepreneurship Education based on Big Data

Lu Bai*, Li Zhu, Yan Li, Siyu Wu, Xuan Yu

Jiangxi Science and Technology Normal University, Nanchang, China

971528034@qq.com

Keywords: Big Data, Innovation and Entrepreneurship Education, Data Analysis Methods and Practical Objectives

Abstract: Innovation and entrepreneurship play an important role in leading industrial development, scientific and technological development, promoting economic growth, creating new employment opportunities, releasing market vitality and building an innovative country. This paper summarizes the IAEE and BDA methods, discusses the current situation of IAEE in CAU, analyzes the training and practical objectives of IAEE, and analyzes the participation of students of different grades in innovation and entrepreneurship projects and competitions. The results showed that the participation of the 1~4 grade college students was on the rise, mainly due to the gradual opening of the entrepreneurship education curriculum for college students, the innovation and entrepreneurship training of Internet plus students and the promotion of challenge cup.

1. Introduction

With the continuous development of computer technology, IAEE, as an educational model, mainly aims at cultivating entrepreneurs and emphasizes the cultivation of innovative personality, including opportunity discovery, entrepreneurial skills, entrepreneurial cognitive development and so on. Through a series of courses or training activities, students can finally master relevant knowledge and business skills and become leaders in business activities.

With the continuous progress of science and technology, many experts have studied IAEE education. For example, Ogochukwu Ij investigated the relationship between entrepreneurship, innovation and finance in Nigeria, and understood the impact of finance on the entrepreneurship and innovation failure of small and medium-sized enterprises in Nigeria. He used Pearson product correlation and standard multilinear regression in SPSS to explore the data to determine the relationship between defined variables and predict the impact of independent variables on dependent variables [1]. Schaeffer P R, Guerrero m, Fischer B explained that the knowledge accumulation of Universities' contribution to the regional ecosystem of innovation and entrepreneurship mainly adopts the view based on the technology transfer process from academia to market [2]. Daub c h, Hasler m, verkuil a h describe an innovative approach to integrating sustainability into business school structures and processes without resistance. It outlines the aspects to be considered and steps to be taken to launch a sustainable entrepreneurship competition to support as many high-quality projects as possible [3]. Although the research results of IAEE are quite fruitful, the research on data analysis methods of IAEE based on big data is still insufficient.

In order to study the data analysis method of IAEE based on big data, this paper studies the IAEE and BDA methods, and finds the detection and evaluation criteria of similar duplicate records. The results show that big data is conducive to the establishment of data analysis methods for IAEE.

2. Method

2.1 Innovation and Entrepreneurship Education

(1) Innovation and entrepreneurship education concept

Innovative education is the behavior of cultivating people with innovative consciousness,

innovative spirit and innovative thinking. Innovative education is not for talented and high IQ students, not for specific groups, but for all students. Cultivating innovative personality is a kind of modern education. Innovative education is not only science, but also innovation in life, education, service, management mode, human resources and so on. The concept of innovative education can be cultivated from an early age, and the habits formed from an early age can be accompanied by a lifetime. Students can teach students in accordance with their aptitude according to their own characteristics, so that students can realize that they are developing and unique people through innovative education. IAEE is the need of the development of the times and an important channel to cultivate talents who meet the needs of the market and society. The market and society are huge and complex, but they are orderly and regular. Therefore, we need such education: through school education and training, they can form a habit of cognition or internalization of knowledge before entering the society and market. IAEE cultivates students' cognition of internalized knowledge through curriculum teaching, holding competitions, lectures, practice and other training methods, that is, observation, insight and understanding in various training methods, so as to bring this cognition into the workplace or society, rather than blindly solving the problems they face. IAEE not only cultivates students' ability to acquire new knowledge in the process of innovation and entrepreneurship, but also cultivates students' spirit of "independent exploration, communication and cooperation". By participating in the corresponding innovation and entrepreneurship competition, students can cultivate students' team spirit together with team members, which is also the quality they need to enter the society and enterprises. Innovation and entrepreneurship emphasizes that students should get rid of the teaching of traditional education, the teaching means of "one speech" and "cramming", the quality education of teachers' guidance and students' autonomous learning, and form their own set of learning methods. This learning attitude and spirit can be used in life, combined with the knowledge learned in school and the experience of entering the society, so as to realize their life value.

(2) Current situation of IAEE in Colleges and Universities

Under the new situation, industry university research innovation and entrepreneurship cooperation has been deepened, and entrepreneurship practice projects have also begun to enter the market. Many local governments also contribute to the IAEE of college students through various channels and methods. First, through the formulation of rent reduction, tax reduction and other preferential policies to promote college students' innovation and entrepreneurship; Secondly, improve the efficiency of College Students' IAEE by coordinating the relationship between enterprises and colleges, such as encouraging enterprises to accept college students' internship through subsidies to enterprises; Thirdly, through the construction of entrepreneurship parks, incubate some good entrepreneurship projects. This shows that the development of IAEE in China has entered the golden stage. We should take advantage of this rare opportunity. In recent years, the Ministry of Education and various provinces and cities have successively issued and promoted the relevant policy documents of "improving the ability of independent innovation and building an innovative country" and "promoting employment through entrepreneurship". Departments at all levels and CAU also actively promote this work by holding or participating in various relevant training and projects, holding academic and practical exchange meetings and other measures.

(3) Training and practical objectives of IAEE

The training and practical objectives of IAEE mainly include the following four aspects: first, awareness training, through learning the basic knowledge of innovation and entrepreneurship and instilling the basic theories such as conceptual elements, so that students have the psychological preparation and skill foundation of innovative and entrepreneurial talents; The second is to improve the ability to decompose the basic ability of innovation and entrepreneurship into different skill steps in an analytical way, that is, from a single observation and analysis ability to a comprehensive critical thinking and organization and coordination ability, so as to enable students to have the necessary entrepreneurial ability. Third, environmental awareness: understand the current enterprise and industry environment from the perspective of business practice and macro perspective, gain insight into tuye information, understand entrepreneurial opportunities, especially correctly

understand and evaluate entrepreneurial risks, understand the characteristics and differences of different business models, and master business skills such as planning and design; Fourth, actual simulation. With the help of the simulation platform, we can truly experience entrepreneurship and company management, specifically participate in the writing of entrepreneurship letters, simulate enterprises and other activities, experience the entrepreneurship of various elements of entrepreneurship, and immerse ourselves in market evaluation, financing and M & A.

2.2 Concept of Big Data Analysis Method

BDA method is one of the scientific research methods different from small data analysis method. Taking massive big data as the analysis starting point, it uses clustering regression analysis, experimental simulation and other methods to peel off the information contained in the data, so as to find new knowledge and present the analysis results in an intuitive form. Focus on in-depth data mining and analysis, and explore the development law of things behind the data. From big data collection to big data processing and BDA, this process always produces new data. At the same time, these data can record the "track" of the analysis process. In addition, the technologies and tools used in the BDA process are intuitive and intuitive, making the analysis process intuitive. The second is the three-dimensional image of BDA results, such as GIS spatial data visualization technology, three-dimensional modeling, scene data interaction, three-dimensional network diagram, three-dimensional or thinking three-dimensional simulation, so as to transform the BDA results into a simpler and intuitive form. BDA methods are different from traditional analysis methods. Intelligent collection and analysis based on full sample data can not only accurately analyze and judge the correlation between different data, but also dynamically warn and scientifically predict the basic trend of things.

2.3 Test and Evaluation Criteria for Similar and Repeated Records

At present, there are many similar duplicate record detection and evaluation criteria, and there are different evaluation criteria for different applications. For the detection and evaluation of similar duplicate records in big data environment, this paper uses memory rate R (recall rate), accuracy P (accuracy) and balance F1 as evaluation criteria. The corresponding calculation method is shown in formula (1):

$$R = \frac{N}{\sum_{i=1}^k N_i} \times 100\% \quad (1)$$

Where n - the number of identified similar duplicate records (pieces), $\sum_{i=1}^k N_i$ - the number of all similar duplicate records (pieces); R - the proportion (%) between the identified similar duplicate records and all similar duplicate records.

IDF: weigh the importance of keywords; The definition of keyword IDF is as follows (2):

$$idf_i = \lg \frac{|D|}{|\{j: t_i \in E_j\}|} \quad (2)$$

Where |D| - the total number of keywords in the data table.

Define TF-IDF to calculate the word frequency weight of keywords, as follows (3):

$$W_{(i)} = IF - IDF(t_{(i)}) = tf_{i,j} \setminus idf_i \quad (3)$$

Where $tf_{i,j}$ represents word frequency and idf_i represents reverse file frequency.

3. Experience

3.1 Object Extraction

The contents related to the practice of IAEE in the public information of university websites should be first-hand materials, and the precipitated contents have high credibility. It should be pointed out that the information about the current situation of IAEE in CAU obtained on the website is difficult to fully reflect the current situation of IAEE in CAU. In fact, getting relevant information from university websites and conducting research have gradually been favored by researchers. For example, in the "undergraduate teaching quality report" published on the official website of the University, many research articles have been published. The sampling method is to take the serial number of the list of 25 ordinary CAU published by the Ministry of Education as the sampling box, and use the randbetween function in Excel table (referring to returning a random number between two specified numbers, i.e. returning random sampling) for random sampling to obtain the random list of 15 ordinary CAU, and finally determine these 15 ordinary CAU as the sampling sample. Conduct targeted interviews, questionnaires, development environment, system functional and non functional requirements investigation and analysis. Interviewed some IAEE experts and excellent teachers in CAU. Taking Internet plus competition as the key point, combining the experience of the innovative and entrepreneurial training program of college students, and forming a teaching mode, focusing on analyzing problems and difficulties, finally forming a questionnaire.

3.2 Experimental Analysis

Main business process of the project: after accumulating certain knowledge, students can obtain competition notice in the information module or independently initiate projects, and initiate team tasks in the team formation module. Other students can choose to sign up. The team leader can query the applicants according to the BDA database to judge whether they agree to join the team. After the successful establishment of the team, the team members will jointly prepare the project plan and enter the project approval stage after passing the teacher's review. After many online reports and guidance, the final results will be reported to participate in the competition or docking business incubation and entrepreneurship projects. First, classify and grade the courses. Through mind mapping, this paper systematically combs all the knowledge systems of PHP project development, which is mainly divided into seven learning parts: Web front-end stage, PHP development foundation, advanced PHP development, server development stage, enterprise development stage, webapp mobile Internet development and project practice stage. Each part contains different learning contents. Secondly, initialize the student preschool education model. Through the online test of three question banks: Web page Foundation, programming foundation and database foundation, combined with the basic personal information stored by students in the platform. Finally, through the analysis of the data model, the judgment of the initial learning ability is given.

4. Discussion

4.1 Analysis of Students of Different Grades Participating in Innovation and Entrepreneurship Projects and Competitions

In terms of innovation and entrepreneurship projects and competitions, we investigated the current situation of teachers and students respectively. By setting up the main problems encountered in innovation and entrepreneurship or competition: team but not creativity, creativity without proper teammates, we analyzed the main pain points of college students in innovation and entrepreneurship projects or competitions, and then investigated the needs of IAEE platform.

Table 1. Chart of students' willingness to participate in innovation and entrepreneurship projects or competitions in different grades

	Yes	No, I want to	No, I don't want to
Freshman	12%	67%	7%
Sophomore	34%	71%	10%
Junior	43%	85%	5%
Senior	85%	92%	15%

It can be seen from the above that 12% of freshmen have participated in innovation and entrepreneurship projects or competitions, 34% of sophomores have participated in innovation and entrepreneurship projects or competitions, 43% of junior students have participated in innovation and entrepreneurship projects or competitions, and 85% of senior students have participated in innovation and entrepreneurship projects or competitions. The specific results are shown in Figure 1.

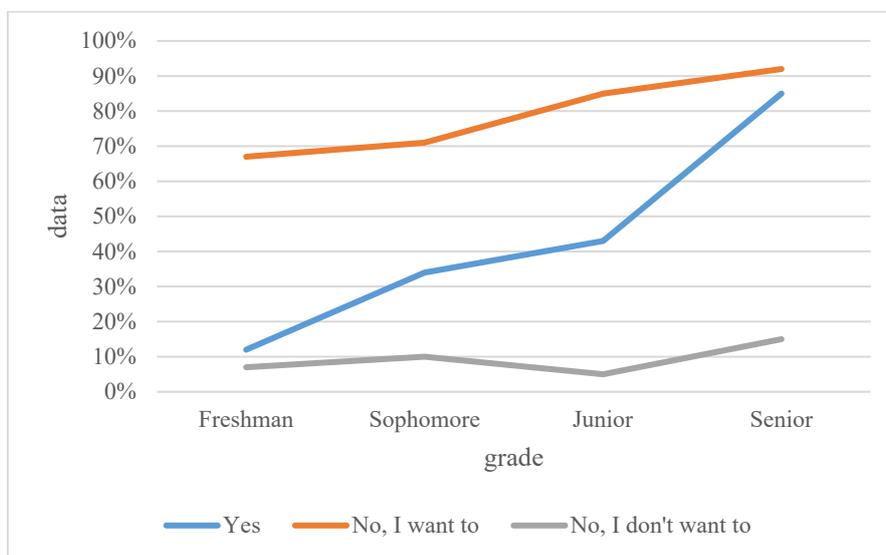


Figure 1. Chart of students' willingness to participate in innovation and entrepreneurship projects or competitions in different grades

From the Internet plus data chart, we can see that the first or two levels rarely participate in innovation or entrepreneurship projects or competitions. They have a strong need to participate in the first grade, and the proportion of participants is 67% and 71% respectively. The number of students from grade one to grade four has been increasing gradually. The main reason is that the innovation and entrepreneurship courses for undergraduates are gradually open, and the Internet plus the challenge cup of innovation and entrepreneurship training for college students is promoted.

4.2 Experimental Analysis

In the verification experiment of accuracy (P) and recall (R), the detection results of SRDA and simhash are compared. The balance between accuracy and recall (F1) was used as the evaluation index of the detection performance of similar duplicate records. The larger the F1 value, the better the detection performance of the algorithm. The test results are shown in Table 2.

Table 2. Detection results of SRDA and simhash algorithms

	P	R	F1
SRDA	0.65	0.82	0.73
simhash	0.84	0.83	0.80

It can be seen from the above that the accuracy rate of SRDA algorithm is 0.65 and that of simhash algorithm is 0.84; The recall rate of SRDA algorithm is 0.82 and that of simhash algorithm is 0.83; The balance between the accuracy rate and recall rate of SRDA algorithm is 0.73, and the

balance between the accuracy rate and recall rate of simhash algorithm is 0.80. The specific presentation results are shown in Figure 2.

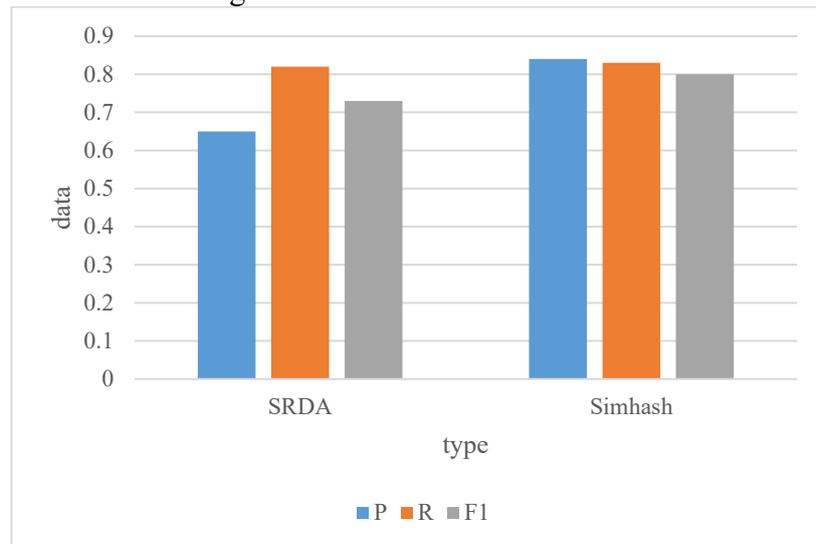


Figure 2. Detection results of SRDA and simhash algorithms

It can be seen from the above that the improved SRDA algorithm has better performance. The P of SRDA is 0.65 and R is 0.82. P and R show a good balance, that is, F1 is 0.73%, showing better detection performance; In addition, the detection accuracy P of SRDA algorithm is about 0.19 lower than simhash algorithm, the recall rate R is about 0.01 lower than simhash algorithm, and the balance F1 is about 0.07 lower than simhash algorithm. Therefore, the improved algorithm SRDA has stronger detection performance and applicability.

5. Conclusion

Innovation and entrepreneurship is a spirit and attitude, and IAEE is a systematic long-term project. The purpose of IAEE is to cultivate students' innovative thinking, entrepreneurial consciousness and entrepreneurial behavior. The core is to cultivate entrepreneurship. Today's society requires CAU to cultivate talents to meet the needs of social development. Therefore, IAEE has become an important entry point for a new round of education and teaching reform. This paper tests the accuracy, recall and balance of SRDA algorithm and simhash algorithm. The results show that the detection accuracy P of SRDA algorithm is about 0.19 lower than simhash algorithm, the recall R is about 0.01 lower than simhash algorithm, and the balance F1 is about 0.07 lower than simhash algorithm. Therefore, the improved algorithm SRDA has stronger detection performance and applicability.

Acknowledgement

Jiangxi Science and Technology Normal University Special Scientific Task (College Ideological and Political Work) Key Project (2020XJGXSZZD005), An Empirical Study on the Cultivation of Non-intellectual Factors to Improve College Students' Innovation and Entrepreneurship Ability.

Jiangxi Science and Technology Normal University Teaching Reform Research General Project (XJJG-20-61-44), Explores the New Teaching Mode of Q-P-M-O of Innovation and Entrepreneurship Education Course in Colleges from the Perspective of Social-embedding.

References

- [1] Ogochukwu I J. Entrepreneurship Innovation and Finance[J]. Journal of Behavioural Economics, Finance, Entrepreneurship, Accounting and Transport, 2021, 9(1):16-35.
- [2] Schaeffer P R, Guerrero M, Fischer B B. Mutualism in ecosystems of innovation and

- entrepreneurship: A bidirectional perspective on universities' linkages[J]. *Journal of Business Research*, 2021, 134(September):184-197.
- [3] Daub C H, Hasler M, Verkuil A H, et al. Universities talk, students walk: promoting innovative sustainability projects[J]. *International Journal of Sustainability in Higher Education*, 2020, 21(1):97-111.
- [4] Rpp A, Mba B, Mn C, et al. The dynamics among entrepreneurship, innovation, and economic growth in the Eurozone countries[J]. *Journal of Policy Modeling*, 2020, 42(5):1106-1122.
- [5] Boone C, Lokshin B, Guenter H, et al. Top management team nationality diversity, corporate entrepreneurship, and innovation in multinational firms[J]. *Strategic Management Journal*, 2019, 40(2):277-302.
- [6] Schuelke-Leech B A. Engineering Entrepreneurship Teaching and Practice in the United States and Canada[J]. *IEEE Transactions on Engineering Management*, 2020, PP(99):1-20.
- [7] Zeng J, Ren J. How does green entrepreneurship affect environmental improvement? Empirical findings from 293 enterprises[J]. *International Entrepreneurship and Management Journal*, 2022, 18(1):409-434.
- [8] Avillez M M D, Greenman A, Marlow S. Ethical Judgments About Social Entrepreneurship in Sub-Saharan Africa: The Influence of Spatio-Cultural Meanings[J]. *Journal of Business Ethics*, 2019, 161(4):877-892.
- [9] Chung D, Kim M J, Kang J. Correction to: Influence of alliance portfolio diversity on innovation performance: the role of internal capabilities of value creation[J]. *Review of Managerial Science*, 2021, 16(3):931-932.
- [10] Xiao Q, Liu H, Wu B. How Bachelorhood and Migration Increase the HIV Transmission Risk Through Commercial Sex in China?[J]. *AIDS and Behavior*, 2019, 24(3):791-801.
- [11] Haltiwanger J. Entrepreneurship in the twenty-first century[J]. *Small Business Economics*, 2021, 58(1):27-40.
- [12] Karlsson C, Rickardsson J, Wincent J. Diversity, innovation and entrepreneurship: where are we and where should we go in future studies?[J]. *Small Business Economics*, 2021, 56(2):759-772.