

Trend of higher education abroad policy analysis based on regression analysis and data mining

Anh Thu Vo, Wanbing Shi

School of Humanities and Law Northeastern University, Shenyang, China

Keywords: Regression analysis, data mining, higher education, study abroad policy, trend study.

Abstract. Educational circles all over the world pay great attention to the study of the internationalization policies and trends of higher education. China higher education policies and regulations continue to develop and improve, in accelerating the education system of governance and governance capacity modernization and the transformation of functions and decentralization, promoting fair education to improve the quality of education that deepen the reform of key areas and achieved new breakthrough. Internationalization of education is the development trend of higher education in the future. Information technology can only promote the reform and development of education with the integration of education. As a mature branch of data mining technology, statistical data mining uses statistical method and data mining principle to analyze and process target data. The principle and application of regression analysis technique are discussed in this paper.

1. Introduction

The United States is the most developed country in the world and the initiator and beneficiary of the internationalization of education in the world today. The United States uses its massive high quality system of higher education and scholarship program to create a multitude of names, attracting the world's young people to come to America to study the environment, in order to keep it in the world's most populous nation's first foreign students. With the arrival of the era of economic globalization and the continuous opening of the international educational market, the scale of the acceptance of foreign students in the United States has continued to expand for many years.

As early as 1946, the U.S. government will set up a "Fulbright Program", to establish the basic principles of American students and scholars of foreign exchange, promote the formation of internationalization of higher education in the United States organized pattern. The program is the flagship of international exchanges in the United states. The aim is to set up scholarships to support American students and scholars abroad, while supporting overseas students and scholars to study and engage in research activities in the United states. In late 1980s the establishment of the American Council on education, an important mission which is the implementation of the Fulbright program.

In 1966, the United States Congress passed the International Education Act, marking international educational exchanges and cooperation as the law of the United States and rising as a national will to educate. This law was the first specialized international educational exchange law, which had a significant impact on all countries in the world. The U.S. government also had been set up to promote the education of foreign students in international education policy mechanism ", responsible for an international educational exchange programs with foreign governments; established by the deputy director of the news department of educational and cultural affairs as chairman, between the government and non-governmental organizations of federal committee, responsible for the improvement of the relationship between the various departments. To improve efficiency, more effective implementation of government funded international exchange and training plan. The U.S. government will study the ultimate goal of education as a national development strategy and policy, and try to make students become a friend of the United States, and America's ideas and values back to their own country, let more people understand the world.

The United States is a decentralized country, and the management rights of higher education belong to the States, but the federal government can exert influence on the international education of schools by means of law and financial appropriation. The United States government supports

American universities and colleges in many ways to recruit international students and provides them with scholarships and grants. For example, the government departments of the Federal Department of education, national student loan supervision organization, the National Science Education Fund and other countries have set up a variety of scholarships and grants for the project, in order to attract and encourage foreign students to apply for American universities, from the economy to foreign students to complete their studies successfully help. In addition, the United States government also supports and welcomes various foundations, social organizations, individuals, as well as bilateral institutions and international organizations, such as the establishment of scholarships, funding for young people studying in the United States.

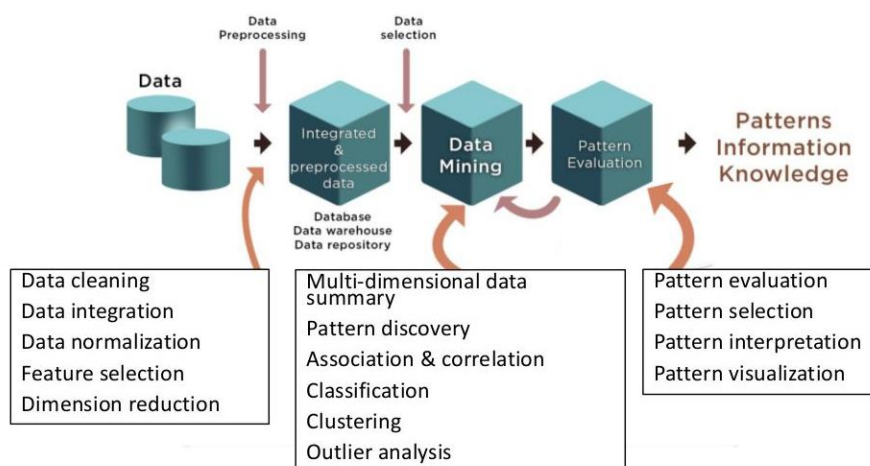


Figure.1 Data mining

2. The Proposed Methodology

2.1 Regression analysis.

Regression analysis is one of the most widely used branches in statistical analysis. It originated from Gauss's least square method in nineteenth Century and formed in the early twentieth Century. Regression is a method of analyzing the relationship between independent variables and dependent variables, which estimates and predicts the total mean of dependent variables based on known independent variables. The mathematical expression obtained by regression analysis is called regression equation. It may be a straight line or a curve. There are many different types of regression in statistics, but their basic idea is that the created model matches the values in the predicted attribute.

With the application of information system based on data warehouse technology, such as database and data warehouse, the large amount of data is continuously produced, resulting in a huge increase in the amount of data stored in the database. However, behind the huge data hiding many important information, such as the information extracted from the database, will be very useful, data mining technology is accompanied by a demand from research to application. Data mining is from a large number of incomplete, noisy, fuzzy, random data to extract implicit, but not yet known, but useful information and knowledge.

Regression analysis is the interpretation of changes in another variable through changes in one or more variables. It includes finding independent variables and dependent variables, setting mathematical models, testing models, estimating and forecasting, etc.. The relationship between variables, there is a definite functional relationship, others did not change with y variables, X variables, but not by the value of X calculated the value of Y, the relationship between X and Y is called relationship. Regression analysis is a mathematical and statistical method to study the correlation between variables.

Statistics is one of the key technologies of data mining, the application is very extensive, there is a relationship between the variables has no definite function relation, but the large amount of observational data, it can be found that there exists a certain statistical regularity between them, study the statistical law or relevant analysis methods to study the relationship between variables is called

regression mathematical statistics. It helps us to estimate the value of another random variable effectively from the value obtained by a variable that can be controlled or can be accurately observed.

2.2 Self-organizing data mining method.

Self-organizing data mining is the core technology of GMDH (group method of data handling), which is an evolution method of evolution and screening based on it, according to a given criterion from the selection of the final model set a series of candidate models. The GMDH algorithm generates many candidate models with increasing complexity through inheritance, mutation, and selection until the model produces a fitting over the observed sample data. The initial model of the method requires a certain amount of the initial model (or neurons) by differential or difference equations, or to describe their solution (e.g., polynomial), in general, should choose the reference function for the object of study characteristics to generate the initial model. Usually higher order Kolmogorov-Gabe polynomials can be used as reference functions.

The GMDH algorithm is characterized by the use of internal and external guidelines for data grouping and throughout the modeling process. It will be observed sample data is divided into training set and test set, the initial model is constructed from the reference function (function) set based on training set using the criterion (least squares) to estimate the parameters to be selected to obtain the intermediate model, intermediate candidate model on the test set by the criterion). By repeating such a process of inheritance, mutation, selection and evolution, the complexity of the intermediate candidate model is increased until an optimal complexity model is obtained.

The regression analysis method determines the model structure (curve type) according to prior knowledge, and then uses the least square method to estimate the coefficients in the whole sample data set. It uses successive regression analysis to filter the variables to produce the optimal regression model. That is, the introduction of the factors into each factor and the introduction of the factor are the sum of the squares of the partial regression, which is significant through the test. At the same time, after each factor is added, the old factor is checked one by one, and the square sum of the partial regression is changed into a non-significant factor. Finally, do the variance analysis and hypothesis test on the final regression model, and determine whether the final regression equation is meaningful or not.

The self-organizing data mining method of data partition, establish the model in the training set, and the use of fresh information (testing data) through the criterion for model selection, it is important to distinguish the characteristics of regression analysis method.

2.3 Continuing and steady study abroad policy.

There are more than 3600 universities in the United States that offer undergraduate degrees, with more than 1700 universities offering a master's degree or above. In order to ensure the quality of international students, American universities have been strict with the entrance. In the aspect of language, for students whose native language is non English speaking countries apply to study in the United States must pass before the language proficiency test, such as: apply for undergraduate degree for foreign students to provide TOEFL scores for students, and the minimum fraction line must reach; apply for master degree students in addition to providing TOEFL results, but also provide GRE or GMAT and LSAT scores. In education, students require bachelor degree required for high school graduates; apply for master's degree must have bachelor degree or above; undergraduate students allowed to have a college degree to students and exchange students to pursue a scientific application etc..

On the one hand, the U.S. government was well aware of the strategic background of globalization, to recruit foreign students for special media, enhance cultural exchanges with other countries, and promote strategic cooperation and common development in the future, the strategic situation of the United States is undoubtedly a wonderful chess. At the same time, foreign students bring ideas, culture and customs of the United States to promote the cultural exchange, bring new ideas and creativity of American society, the social culture more all-embracing.

On the other hand, in recent years, due to the economic crisis, the U. S. economy in the world's position has suffered a certain degree of blow. Enrolling a large number of foreign students is a considerable source of economic growth for education resources in the United States and even

government spending. According to the US Department of Commerce, in 2011, foreign students and their families studying in the United States and living expenses contributed \$20 billion 200 million to the U.S. economy, creating at least 105 thousand jobs. Students with their knowledge, expertise and diversification of the idea that American campus and society more rich and varied, and directly promoted the economic development of the United States, has brought huge economic benefits to the United states. To this end, the United States government in the implementation of the stability of the study abroad policy, and actively take measures to further relax the policy of studying abroad, in order to ensure the continued growth of the total size of students.

There are many factors that cause China students study abroad, in addition to the United States foreign policy easing the environment, employment market competition superiority, economic condition is fierce, more and more parents to middle or high school at home and children to the United States, they hope to accept high quality education as early as possible, as soon as possible into the overseas campus life. The younger generation is also an important reason. According to reports, in 2011 the candidates for TOEFL candidates in China, the proportion of candidates under 18 years of age increased by more than 1 times. The proportion of students studying abroad accounts for 8% of the total number of students studying abroad. This data is bound to have a higher growth rate with the development of high school study market and the overseas enrollment system. The expanding size of young study abroad indicates that the scale of future undergraduate study will be expanded.

Some experts predict that in business, engineering, art and design are still the key China student choice, but with the development of international cultural exchanges have become increasingly frequent, the energy, environment and other aspects concerned to strengthen, so the energy and environmental engineering, law, education, journalism, social work and other professional will be more forward-looking. In addition, the United States immigration enforcement bureau expanded to extend the internship period of foreign students in 2011, in addition to the four areas of science, technology, engineering and mathematics, has added a number of areas, involving 50 professional graduates in the United States, and selective internship time can be up to 29 months.

3. Conclusion

In the current embassy management of overseas students focus on cultivating highly level and talents introduction and at the same time, for such a large number of energetic, enthusiastic and thoughtful young people to explore a set of innovative features, methods of work, guidance and strong support China student association work well, let the students learn exotic, campus life pleasant, consciously grow into a love of the motherland, the heart of the country, study hard, excellent students aspiring, for the future of China's modernization, China people's friendship and exchanges and cooperation to contribute. Statistics is one of the key technologies in data mining, the application is very extensive, there is a relationship between the variables has no definite function relation, but the large amount of observational data, it can be found that there exists a certain statistical regularity between them in mathematical statistics to study these statistical law or the relevant analysis methods to study the relationship between variables is the so-called return.

Acknowledgments

Fund Project: National Fund Pedagogy Project of Social Sciences "The Evaluation System Research of Universities' Scientific Research Performance of Humanistic and Social Science" (Project No.: BFA150043)

References

- [1] Witten, Ian H., et al. *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, 2016.
- [2] Alcala-Fdez, Jesus, et al. "Comparison of KEEL versus open source Data Mining tools: Knime and Weka software." (2016).

- [3] Wang, H. and Wang, J., 2014, November. An effective image representation method using kernel classification. In *Tools with Artificial Intelligence (ICTAI)*, 2014 IEEE 26th International Conference on (pp. 853-858). IEEE.
- [4] Alcala-Fdez, Jesus, et al. "Comparison of KEEL versus open source Data Mining tools: Knime and Weka software." (2016).