ICIC Express Letters
Part B: Applications
Volume **11**, Number **2**, February **2020**

ICIC International ©2020 ISSN 2185-2766
pp. **137–144**

EXPLORING THE PERFORMANCE OF SELECTED DOUBLE FIRST-CLASS UNIVERSITIES IN CHINA BY USING IPO MODEL

Ying Xia[1,*] and Dian-Fu Chang[2]

[1]Graduate Institute for Taiwan Studies
Xiamen University
No. 422, Siming South Road, Siming District, Xiamen 361005, P. R. China
*Corresponding author: xiaying0520@foxmail.com

[2]Graduate Institute of Educational Policy and Leadership
Tamkang University
No. 151, Yingzhuan Road, Tamsui District, New Taipei City 25137, Taiwan
140626@mail.tku.edu.tw

Received August 2019; accepted November 2019

Abstract. *The study aims to tackle institute performance and its impact factors related to the double first-class universities in China. To achieve this purpose, this study develops an IPO (input, process and outcome) indictor system to collect data and build test models. There are 10 critical indicators in the proposal model which consists of input, process and output dimensions. This study conducted cluster analysis and regression analysis to disclose the data set. The findings reveal there are three significant groups in the selected double-first class universities. The regression models suggest numbers of senior professors and sufficient budget are crucial factors to interpret the output among these universities. The findings may provide useful information for related policy makers.*
Keywords: Cluster analysis, Double first-class universities, Higher education, Regression analysis

1. **Introduction.** World-class universities movement has been on the way and impacts on higher education system in various countries. In recent years, numerous countries have been providing special funds for university development. For example, the supported fund from UK University Grants Committee, the United States Department of Education with the project fund, Japan and Germany also provided fund for enhancing universities [1-3]. Under such circumstances, China is not an exception. She has been providing funds for a number of major universities to chase the game. In September 2017, the Chinese Ministry of Education, Ministry of Finance and National Development and Reform Commission released the detailed lists of universities and disciplines to be developed under China's double first-class initiative. The double first-class initiative is China's largest education development scheme to date, aimed at increasing the global recognition of China's university system by 2049 (the 100-year anniversary of the establishment of the People's Republic of China). The plan was first announced in November 2015 and the implementation plan was issued in January 2017 [4].

Under this initiative, 42 universities have been identified as having the potential to develop as world class, including 36 universities categorized as type A (already well on the way to being world class) and 6 universities as type B (considered to have the potential to be world class). In addition, 465 disciplines from 140 universities (including the group of 42) are also identified as having the potential to become world class. The list of 42 universities includes all 39 of the former "985" project universities, plus three additional

DOI: 10.24507/icicelb.11.02.137

universities from the former "211" project: Zhengzhou University (Henan), Yunnan University (Yunnan) and Xinjiang University (Xinjiang). The discipline development list is composed mainly of former "985" and "211" project universities, with an additional 25 none "211" project universities that have strengths in particular fields [4]. According to the document announcement, a number of reforms invent to Chinese universities including:

- Strengthen and improve Party leadership at the individual university level and improve the internal governance structures;
- Improve performance management;
- Accelerate talent training and collaborative science education;
- Build community participation and improve links with industry and research institutes and;
- Promote international exchanges and cooperation with world-class universities and academic institutions, to strengthen international collaboration and attract foreign students.

Previous studies have shown different perspectives to address the double first-class universities in China. For example, from the structural viewpoint, the selected universities focusing on science and technology (49%) in terms of humanity and social science only 11% and 18% respectively, the other 22% belongs to medicine and agriculture [5]. Specifically, there are 162 schools focusing on natural science and the other 77 schools belong to social science, art, and education among the selected disciplines [6]. In addition, how to balance national goal and universities' strategies is important. It is not necessary that the universities should neglect the traditional mission and target with themselves in the process of chasing current ambitious goal toward world-class universities [7]. Moreover, the selected universities were not quite located in the disadvantaged provinces. It is another worrisome phenomenon in current policy implementation which did not fit the equality principle [8]. Analyzing the selected indicators, related studies found there is a negative relationship between the selected schools and number of undergraduate students among the universities. While the number of master and doctoral students has made differences to support the listed universities. The academic publications in WOS (web of science) has become a crucial element within the double-class universities. The studies found the funding, senior professor, and number of elite scholars also support the selected universities [9]. While using systematic models to tackle the issues among the double first-class universities is very limited in current literature.

This study targets to develop a framework to realize the meanings of the investment for the selected universities. Based on the research focusing, this paper selected the 41 so called double first-class universities (excluding one military related university) as the research target. According to the data in 2017, this study tackles the similarity or dissimilarity of the selected universities to achieve their goal of world class universities. With this regard, this study aims to explore the equality of these selected 41 universities by using IPO (input, process, and output) models. Given this purpose, the research questions are listed as follows.

a. Do the selected double first-class universities have similarity on current performance?
b. What kind of factors may impact on the outcome of current selected universities?
c. Which model can be used to interpret the phenomenon?

The result will display the current similarity and dissimilarity performance among the first-class universities. The main factors which impact on the outcome of the universities will be selected. Finally, the regression models will be built to interpret the phenomenon. The structure of this paper begins with the method section, which will address how the similarity of the universities has been detected and how the models have been selected.

Then, the result will display based on the related evidences from data transformation. Finally, the conclusion will draw based on the findings.

2. Method.

2.1. **Definition of IPO model.** The CIPO (context, input, process, and outcome) has been considered as a full model to interpret a systematic phenomenon [10]. While IPO can be as an alternative approach to dealing with the systematic data when context data are not available. In this study, the IPO model consists of three dimensions, namely input, process, and output dimensions. Since the data of context (C) indicators are not easy to collect, the context indicators will not include in the model. For data collection's reason, it is our crucial concern. Moreover, we selected the indicators to fit the model based on previous literature and their importance. Totally, we selected 10 key indicators to represent the main content of IPO to fit the study's purpose for double-class universities. The input dimension includes expenditure (I1), full-time faculty (I2), senior professors (I3) and senior professor ratio (I4). The process dimension includes student numbers (P1), undergraduate vs. graduate ratio (P2), and international students (P3). The output dimension includes graduate employability (O1), hi-impact articles (O2) and SCI & SSCI articles (O3). In the data set, the expenditure indicator has found five missing values, including Renmin University of China, Beihang University, Harbin Institute of Technology, Northwestern Polytechnical University, and Yunan University.

2.2. **Data collection.** This study conducted cluster analysis, regression model to explore the IPO data among 41 double-first class universities in China, which does not include National University of Defense Technology. The target data were based on the following sources:

a. The Website of Ministry of Education of the People's Republic of China [11];
b. The Report on Undergraduate Teaching Quality, Taking the Peking University for Example [12]. Most of the reports are from 2017 to 2018, the other 9 of them are from 2016 to 2017, including Renmin University of China, Beihang University, Nankai University, Ocean University of China, Huazhong University of Science and Technology, Sichuan University, Xi'an Jiaotong University, Northwestern Polytechnical University and Lanzhou University;
c. The world class ranking data set done by Shanghai Jiaotong University [13].

In this study, we applied Minitab package to conduct related interpret models.

2.3. **Cluster analysis for group difference.** Cluster analysis was used to determine the optimal number of clusters to interpret the selected data set by using different evaluation criteria. We conduct the cluster analysis following the steps: First, the data based on the performance purposes were selected in terms of two output indicators included in the model; Then, we consider which model fit the data set, K-means or hierarchical clustering will run by Minitab package to determine the fittest clusters. Basically, the cluster algorithms are as follows:

- Select k point as initial centroids,
- Repeat,
- From k clusters by assigning each point to its closest centroids (this section refers to the distance transformation),
- Re-compute the centroids of each cluster,
- Until centroids do not change.

To get fitted models, this process may try couple times with different methods in cluster analysis. The dendrogram function plots the cluster tree. Based on the dendrogram, the result will display both the cluster and sub-cluster relationships [14-16]. A dendrogram consists of many U-shaped lines that connect data points in a hierarchical tree. The

height of each U represents the distance between the two data points being connected. If there are 30 or fewer data points in the original data set, then each leaf in the dendrogram corresponds to one data point [17].

2.4. **Regression analysis for better interpreting models.** In this study, we considered the output indicators as the responsive variables in the regression models. We assumed there are three regression models existing in the output dimension. We selected stepwise method to build fittest models. Stepwise regression is a combination of the forward and backward selection techniques. Even though this method is controversial in data mining, it is a convenient approach to determine the major impact factors in the model [18]. Stepwise regression is a modification of the forward selection so that after each step in which a variable was added, all candidate variables in the model are checked to see if their significance has been reduced below the specified tolerance level. Stepwise regression requires two significance levels: one for adding variables and one for removing variables. The cutoff probability for adding variables should be less than the cutoff probability for removing variables so that the procedure does not get into an infinite loop [19]. In this study, the stepwise selection followed the criteria that α to enter = 0.05, α to remove = 0.05. The Minitab will display the analysis of variance, model summary, coefficients, and regression equation.

3. **Result.** In this section, we present the basic statistics of the selected key indicators. In the following, we address the result of cluster analysis. Finally, the regression models will be displayed. The statistics of the key indicators are listed as Table 1.

TABLE 1. The main features of the selected indicators

Variables	Total	Mean	St. Dev	Minimum	Q1	Maximum	Skewness
expenditure (I1)	41	58.49	36.34	9.10	36.40	193.35	1.95
senior professors (I3)	41	1944	729	641	1461	3613	0.45
undergraduate vs. graduate ratio (P2)	41	1.381	0.673	0.425	0.874	3.558	1.54
employability (O1)	41	0.95799	0.02452	0.88010	0.94975	0.99030	−1.20
hi-impact articles (O2)	41	386.9	298.8	23.0	155.0	1373.0	1.42
SCI & SSCI articles (O3)	41	18698	10904	808	11397	44674	0.70

3.1. **The findings of cluster analysis.** The result of cluster analysis shows there are three clusters with performance data in the output dimension (O1 and O2). The final partition reveals Cluster 1 with 10 universities, Cluster 2 with 12 universities, and Cluster 3 with 19 universities, see Figure 1. With ward linkage and Euclidean distance, we found the cluster analysis works well. Based on the information of clusters, we can find the Cluster 1 is listed first, then Cluster 2 and Cluster 3 with their percent of employability and numbers of high-impact articles. The details of cluster centroids are listed in Table 2.

3.2. **Selection of regression models.** The regression models were built by using stepwise method with specific output indicators as dependent variables. We found, the employability (O1) can only be explained by the undergraduate vs. graduate ratio (P2). The R^2 is .376 in this model, it implies the employability (O1) among the 41 universities can be explained 36% (adjusted) by the undergraduate vs. graduate ratio (P2). The fitted line plot is displayed in Figure 2. The result reveals O1 and P2 with negative relationship. It means the higher undergraduate students the lower employability among these universities. Table 3 demonstrates the analysis of variance and coefficient in the model.

The regression model is listed as follows.

$$\text{employability (O1)} = 0.9889 - 0.02236 \text{ undergraduate vs. graduate ratio (P2)}$$

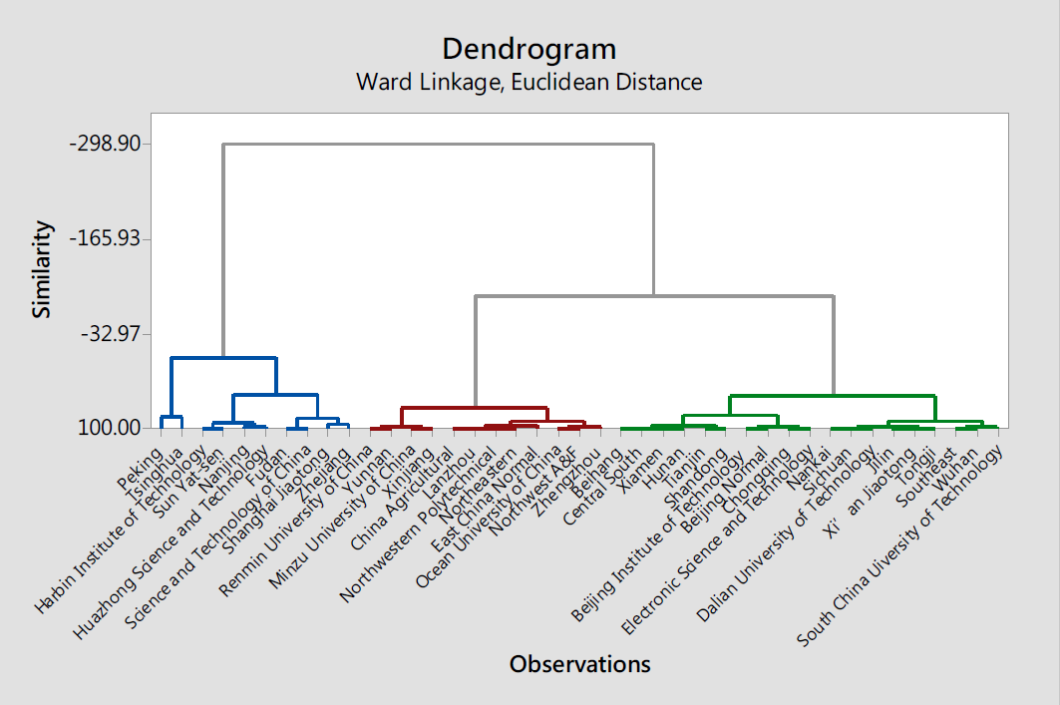

FIGURE 1. The dendrogram with similarity of output among the 41 universities

TABLE 2. The information of cluster centroids

Variable	Cluster 1	Cluster 2	Cluster 3	Grand centroid
employability (O1)	0.967	0.939	0.966	0.958
hi-impact articles (O2)	814.400	104.250	340.421	386.902

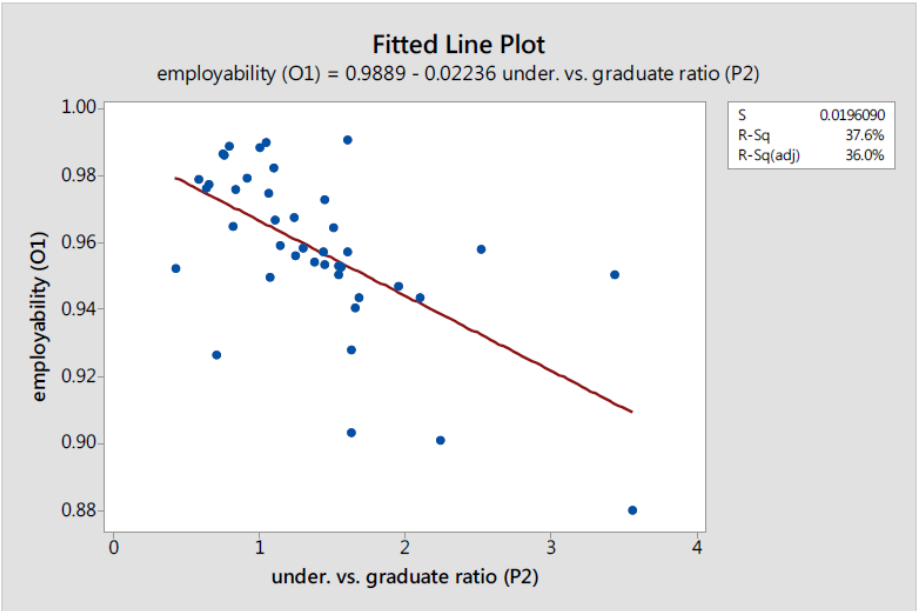

FIGURE 2. Fitted line plot for employability (O1) and P2

TABLE 3. Analysis of variance and coefficient for O1 model

Source	**DF**	**Adj SS**	**Adj MS**	***F*-value**	***p*-value**
Regression	1	0.008527	0.008527	20.35	0.000
undergraduate vs. graduate ratio (P2)	1	0.008527	0.008527	20.35	0.000
Error	34	0.014244	0.000419		
Total	35	0.022771			
Term	**Coef**	**SE Coef**	***t*-value**	***p*-value**	**VIF**
constant	0.9889	0.00765	129.18	0.000	
undergraduate vs. graduate ratio (P2)	−0.02236	0.00490	−4.51	0.000	1.00

The O2 indicator refers to the numbers of high-impact articles. The result reveals the higher expenditure support by government, the higher hi-impact articles might produce from the universities. The interpret regression model is listed as follows.

$$\text{hi-impact articles (O2)} = -38.67 + 7.569 \text{ expenditure (I1)}$$

This study displays the related regression model information in Table 4. The result includes the analysis of variance for the regression model with sum of square (SS) by F-test and t-test. The expenditure (I1) has become the largest contributor for hi-impact articles (O2). The equation shows both variables can be explained 78.6% of variances. The relationship between predictor and response variable was presented as Figure 3.

TABLE 4. Analysis of variance and coefficient for O2 model

Source	**DF**	**Adj SS**	**Adj MS**	***F*-value**	***p*-value**
Regression	1	2647827	2647827	129.89	0.000
expenditure (I1)	1	2647827	2647827	129.89	0.000
Error	34	693117	20386		
Total	35	3340944			
Term	**Coef**	**SE Coef**	***t*-value**	***p*-value**	**VIF**
constant	−38.67	45.6	−0.85	0.402	
expenditure (I1)	7.569	0.664	11.40	0.000	1.00

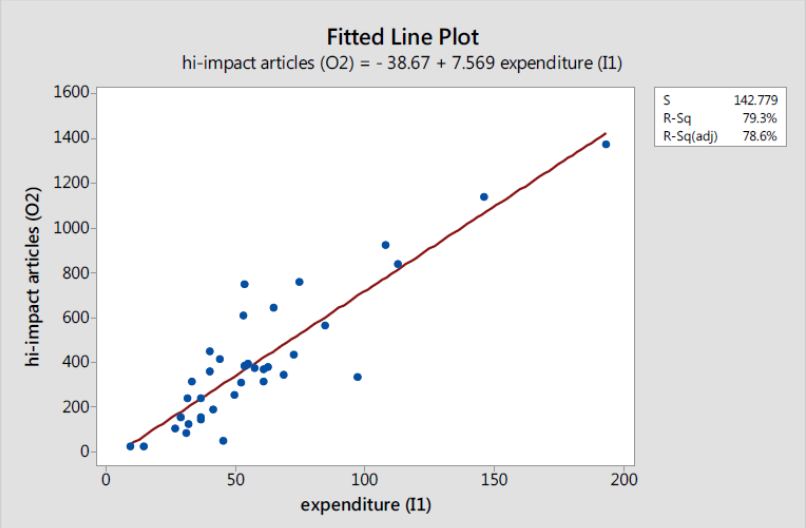

FIGURE 3. Fitted line plot for hi-impact articles (O2) with I1

The SCI & SSCI articles (O3) indicator refers to the journal papers in SCI and SSCI collected by Thomson Reuters. In the regression model, we found expenditure (I1) and senior professors (I3) are main contributors for the production of international journal papers. The model can be explained by 78.48% of variances among them. The regression equation can be written as

$$\text{SCI \& SSCI articles (O3)} = -4034 + 168.3 \text{ expenditure (I1)} + 6.69 \text{ senior professors (I3)}$$

The related tests for the regression model are listed as Table 5. Surface plot of O3 with I1 and I3 is demonstrated in Figure 4. The results reveal the sufficient funding supported with senior professors can make differences the knowledge production in Chinese higher education setting.

TABLE 5. Analysis of variance and coefficient for O3 model

Source	**DF**	**Adj SS**	**Adj MS**	***F*-value**	***p*-value**
Regression	2	3519389377	1759694688	64.83	0.000
expenditure (I1)	1	781152333	781152333	28.78	0.000
senior professors (I3)	1	514461966	514461966	18.95	0.000
Error	33	895751733	27143992		
Total	35	4415141110			
Term	**Coef**	**SE Coef**	***t*-value**	***p*-vlue**	**VIF**
constant	−4034	2531	−1.59	0.121	
expenditure (I1)	168.3	31.4	5.36	0.000	1.68
senior professors (I3)	6.69	1.54	4.35	0.000	1.68

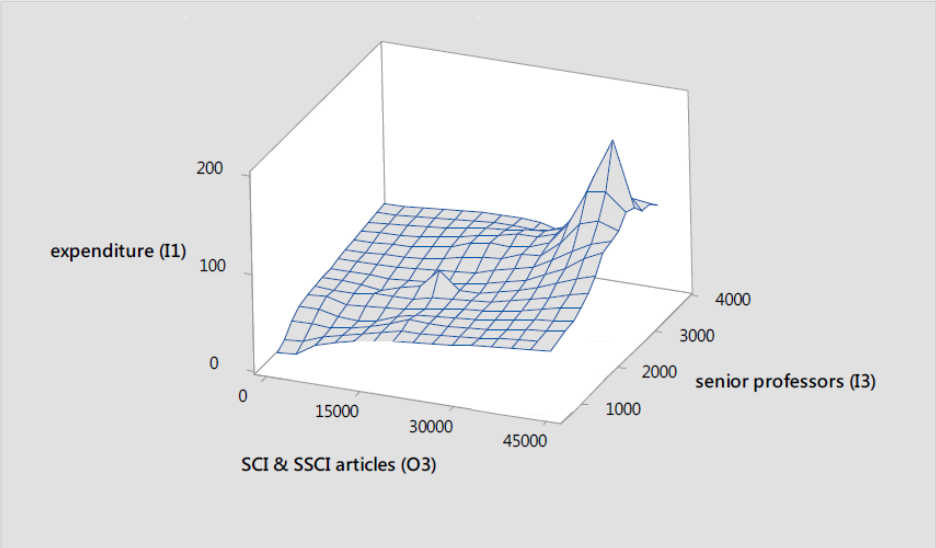

FIGURE 4. Surface plot of O3 with I1 and I3

4. **Conclusion.** The global competition in higher education is higher than any previous times. The actions for building world-class universities did not stop automatically in any specific developed country. China as one of crucial economies in the world, she will not be absent in the world-class competitive movement. Therefore, how to review the project properly and effectively has become a crucial component in the process of competition. This study demonstrates how the data can be collected by meaningful purposes and how the result can be interpreted by fitted methods. The findings reveal the funding is the crucial component for developing world-class universities under current

initiatives. In addition, the number of senior professors also plays an import role for the selected universities in this process of competition. The proposal IPO framework can be used to achieve this purpose to tackle the issue within the selected 41 universities. The cluster analysis provides the related information with three significant groups which can be used for policy purposes. This study provides an example which also can be extended to tackle the similar issue in other higher education settings. For further studies, we suggest adjusting the related indicator for specific higher education systems which will enhance the capabilities of their problem solving. Moreover, dynamic data linkage system is important for prompting data transformation. We suggest thinking about building a dynamic data linkage system which can enhance the related data collection immediately from universities' web pages. It may need more related experts to engage to fulfill the challenging task.

REFERENCES

[1] J. Salmi, *The Challenge of Establishing World-Class University*, The World Bank, Washington, DC, 2009.

[2] J. C. Shin, Building world-class research university: The brain Korea 21 project, *High Education*, vol.58, pp.669-688, 2009.

[3] D.-F. Chang, C.-T. Wu, G.-S. Ching and C.-W. Tang, An evaluation of the dynamics of the plan to develop first-class universities and top-level research centers in Taiwan, *Asia Pacific Educational Review*, vol.10, no.1, pp.47-57, 2009.

[4] Ministry of Education of the People's Republic of China, *The List of the Double-First Class*, http://www.moe.gov.cn/s78/A22/A22_ztzl/ztzl_tjsylpt/sylpt_jsgx/201712/t20171206_320667.html, 2017.

[5] F. Huang, Double world-class project has more ambitious aims, *University World News*, https://www.universityworldnews.com/post.php?story=2017092913334471, 2017.

[6] G. Newton-Tanzer, How will China's double first class project change higher education in China?, *HUFFPOST*, https://www.huffpost.com/entry/how-will-chinas-double-first-class-project-change_b_5a20200fe4b02edd56c6d748, 2017.

[7] D. Bie, On construction of double first class universities and academic disciplines, *China Higher Education Research*, no.11, pp.7-11, 2017.

[8] C. A. Gao, Closer look at China's world class universities project, *The Diplomat*, https://thediplomat.com/2017/09/a-closer-look-at-chinas-world-class-universities-project/, 2017.

[9] J. Shi, Similarity in form and similarity in spirit: Constructing world-class universities with Chinese characteristics, *China Higher Education Research*, no.3, pp.8-12, 2018.

[10] D.-F. Chang and N.-J. Lin, Applying CIPO indicators to examine internationalization in higher education institutions in Taiwan, *International Journal of Educational Development*, vol.63, pp.20-28, 2018.

[11] Ministry of Education of the People's Republic of China, *Ministry of Education's Departmental Budget for 2017*, http://www.moe.gov.cn/srcsite/A05/s7499/201704/t20170407_302152.html, 2017.

[12] Information Open Network of Peking University, *Report on Undergraduate Teaching Quality of Peking University*, http://dean.pku.edu.cn/notice/content.php?mc=61313&id=1515631702, 2017.

[13] Shanghai Ranking, *Best Chinese Universities*, http://www.zuihaodaxue.com/zuihaodaxuepaiming2017.html, 2017.

[14] P. Kaur, M. Singh and G. S. Josan, Classification and prediction based data mining algorithms to predict slow learners in education sector, *Procedia Computer Science*, vol.57, pp.500-508, 2015.

[15] M. J. Brusco, R. Singh, J. D. Cradit and D. Steinley, Cluster analysis in empirical OM research: Survey and recommendations, *International Journal of Operations & Production Management*, vol.37, no.3, pp.300-320, 2017.

[16] D.-F. Chang and C. C. Chen, Cluster analysis for student performance in PISA2015 among OECD economies, *ICIC Express Letters, Part B: Applications,* vol.9, no.11, pp.1139-1146, 2018.

[17] MathWoks, *Dendrogram*, https://www.mathworks.com/help/stats/dendrogram.html, 2019.

[18] P. Flom, *Stopping Stepwise: Why Stepwise Selection is Bad and What You Should Use Instead*, https:// towardsdatascience.com/stopping-stepwise-why-stepwise-selection-is-bad-and-what-you-should-use-instead-90818b3f52df, 2019.

[19] NCSS Statistical Software, *Chapter 311 Stepwise Regression*, https://ncss-wpengine.netdna-ssl.com/wp-content/themes/ncss/pdf/Procedures/NCSS/Stepwise_Regression.pdf, 2019.