

Statistical Prediction and Marketing Recommendation of Foreign International Students' Consumer Behavior

Xinting Shen*

Simon Fraser University Vancouver, Canada

*Corresponding author: xinting_shen@sfu.ca

Abstract. Under a dynamic and complex educational market, it forces shape the educational environment. In the context of China, accelerating economic growth produces multiple newly Chinese multinational education institutions, which lack accurate analysis of consumer preference's inherent characteristics with educational needs. Therefore, this research is vital in helping new Chinese multinational education institutions make decisions based on foreign countries' students' consumer preference and further filling the Chinese multinational education institution preference analysis' gap. In statistics, this paper uses the Data collected from OECD/UIS/Eurostat (2021) Table B6.1, throughout 45 countries, ranging including bachelor's degree, master and doctorates foreign countries students studying in China, to conduct regression analysis intensely observing foreign international students' Country of Attendance preference. In Marketing, Multi-factor integration model authenticates the overall international student's consumer performance. It is proved that Chinese educational institutions' attraction is dominantly attributed to stable economic growth, advanced information, and communication technology. Specifically, China has a higher affinity towards OECD country students for courses of tertiary, bachelor, master, and doctoral studies. Foreign international students' preference statistics prediction improved the accuracy of foreign international students' behaviors towards the Chinese educational area, driving Chinese educational institutions to a more precise and effective marketing strategy. These results shed light on foreign international students' preference for Chinese education, and how should educational institutions change their marketing methods next.

Keywords: Cross-border; Consumer behaviors research; regression model; Integrated factor analysis model.

1. Introduction

Cross-border education is the education acquired outside students' home country, and China is among the leading countries in receiving international students. According to Statista (2022) [1], there has been an increase in Chinese international students who study abroad. Especially in the United States, Chinese students have increased from 157,558 in 2010 to 317,299 in 2021. However, the statistics of international Students Introduced to China has seldom been analyzed by the statistical analyst. As the most significant population country with stable economic growth, attracting international students improving Chinese citizens' critical thinking competency and cultural Inclusivity, help furtherly develop China's economy. Therefore, analyzing the international student's preference who aimed to study in China is of great importance, and the related research gap urgently needs to be filled.

Predecessors have conducted considerable research based on educational institution marketing and statistical prediction contexts. According to statistical consumer preference' behavior prediction research, reasonable marketing research and statistical analysis make consumer qualitative and quantitative criteria possible to find, which establishes the foundation's meaningfulness for all educational studies [2].

According to predecessors' static marketing research, consumer behavior motivators have also been interpreted into predictable theory, e.g., Planned behavior theory [3], which indicates most people make rational behavior based on complex psychological processes. Brand attachment feasible analysis indicates that brand attachment impact plays a preponderant role in selecting educational institutions [2]. The structural equation model also analyzes consumer behavior upon K12

educational transformation. Structural equation model (SEM) [4]. To adopt to the factor analysis, the main interpretation methods are Exploratory factor analysis and confirmatory factor analysis to deeply analyzed the relationship between variables and other latent factors.

To explore the foreign international student's consumer preference concerning Chinese educational institutions. This research conducts from two dimensions: dynamic analysis and static analysis. To be specific, in static analysis, an integrated multifactor analysis model is constructed to analyze the unobserving latent factors' effects. The cognitive level and behavior attitude, to identify potential Chinese education service preference dimension. In dynamic analysis, the regression analysis will be used to process the observing data based on 45-countries foreign international students who study in China boundary ranging from bachelor to doctorates.

Single theory analysis often experiences cyclicity and incompleteness in performance in static analysis. Therefore an integrated model (multifactor approach) is adopted for static analysis to optimize this issue. The integrated model is established based on the Concept of diversification by analyses of multiple weighted consumer preferences to soften negative foreign countries' consumer' effects on Chinese education institutions to increase Chinese international education institutions' outperformance [5]. Performing an integrated consumer preference model is able to solve a puzzle; the players need the right pieces and the correct combination to achieve the desired outcomes [6]. The mainly used marketing theories are demonstrated as follows.

1.1 Planned behavior theory (TPB theory)

Planned behavior theory (TPB theory) indicates that consumer intention is immediately decided by consumer behavior. As shown in Fig. 1, Three dominant tributes form Intention: behavior attitudes, subjective norms, and behavior control. Behavior attitude is stimulated by behavioral beliefs, which include evaluation of outcomes and beliefs that behavior leads to specific outcomes.

Subjective norm is stimulated by normative beliefs, including beliefs of important referents and motivation to comply with referents. The extended theory established on Planned behavior theory (TPB theory) is the reasoned action theory (TRA theory), developed by Ajzen [7]. This theory assumes all individuals are rational in making the decision, and all behavior needs a complicated psychological process to conduct. In reality, the reasoned action theory cannot be completely conducted because not all behavior is under volitional control. Subconsciousness also plays tremendous importance in consumer decisions.

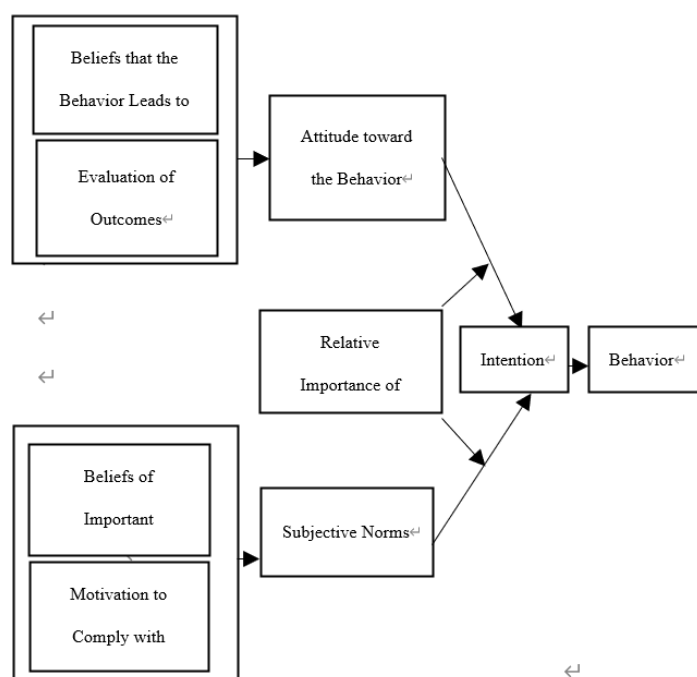


Fig 1. Planned behavior theory (TPB theory) Illustration.

1.2 The Psychoanalytic Theory: Desires & Emotions

The Psychoanalytic Theory indicates that multiple subconscious motivations stimulate consumer behavior, and the consumer cannot fully understand their behavior [8]. The behavior motivator is beyond income, age, and country, hugely influenced by emotions, aspirations, fears, feelings, and desires.

In dynamic analysis, based on observing Data in OECD/UIS/Eurostat (i.e., Table B6.1 database), statistical analysis is conducted to illustrate the relationship and preference of foreign international students [9].

1.3 Correlation coefficient table

Based on 45 countries' foreign international students' data who study in China, the correlation coefficient table has provided noticeable results, demonstrating a significant Pearson correlation between all tertiary courses and bachelor's, master's, and doctorate courses in China. Pearson correlation coefficient (PCC) is a covariance's normalized measurement approach [10]. The ratio of variables covariance and standard deviations product, which reflected linear correlation. The Pearson value is inside the $[-1, 1]$ interval. The sample value is expected to be significantly greater than 0 but less than 1, because 1 represents the unrealistically perfect result, which is nearly impossible to happen in an unperfected world. This dynamic analysis analyzed the relationship between all tertiary education and the bachelor's, master's, and doctorate courses in China, ranging from foreign international students. The coefficient table result indicates that students from all works of life and levels find China an exciting country to learn.

1.4 Goodness-of-Fit estimation

In the re-validation process of the model, the tested indicators include R, R square, Adjusted R square, Standard error of the estimate, and changed statistics. R indicates the correlation between observed values and model-simulated predicted values, and R square indicates the proportion of the response variance explained by predictors. The Goodness-of-Fit is the degree of fit of the regression line to the observed values. The maximum value of R^2 is 1. The closer the value of R^2 is to 1, the better the regression line fits the observed values. Conversely, the smaller the value of R^2 is, the worse the fit of the regression line to the observed values. The standard error of the estimate (SEE) indicates the standard variance of the model's error component, which is also the square root of the squared mean of the residuals in the ANOVA variance table. The smaller the Standard error of the estimate (SEE) is, the better Goodness-of-Fit it will be. DF indicates the optional degree of freedom. The degree of freedom of regression analysis is related to the number of estimated coefficients (Sample size minus restricted conditions). The SIG-F value is the proportional relationship or ratio between the square mean of (regression) and the square mean of (residual). In SPSS variance analysis, the size of SIG-F has no special significance for the degree of the fitting. However, SIG can represent the significance level, i.e., $SIG < 0.05$ is significant in the test. Therefore, the absolute value of the regression coefficient is significantly greater than 0, and the independent variable can predict the dependent variable with 95% confidence that the conclusion is correct.

The rest part of the paper is organized as follows. The Sec. II will be specific Methodology for Statistical analysis of foreign international student's educational preference. The Sec. III will be results based on data analysis and discussion of specific marketing strategy recommendations for the foreign international educational preference in China based on data analysis results. The limitations of this study are also concerned in this part, presenting a specific outlook for future research. Eventually, a brief summary will be given in Sec. IV.

2. Methodology

This paper utilizes regression analysis and ANOVA test statistical method. The data was mainly collected from OECD/UIS/Eurostat, Table B6.1. to help understand the consumer behavior of cross-

border education in China [5]. It contains details of 45 countries globally that have their students in China. The students from other countries studying in China assume different levels of education, ranging in all categories, including bachelor's degrees, master and doctorates. The sample size is $n=45$ countries.

Independent variables are defined as variables that do not influence the outcome of the independent variable. The independent variable is education, while the dependent variables are the consumer behaviors per country. Education is an independent variable since it remains constant across countries. However, different factors will influence the outcome of educational competence within the institution. Research shows that consumer behavior is at par with the institution of higher learning, and the institution will attract students from the home country and vice versa. *Dependent variables* are defined as variables that influence the outcome of the independent variable. In this case, education is the independent variable. The factors that could affect its outcome include technology advancements, consumer behavior, Institutional brand marketing, human resources, and resource allocations. After introducing free primary and secondary education for its citizens, China became a privileged country, thus improving the quality of education and increasing its attractiveness to the outside world.

For the analysis process, Linear regression analysis is a main method of predicting the future value of a random variable correlated with one or a group of independent variables based on the variation of the variable. Regression analysis requires the development of a regression equation that describes the correlation between the variables. Depending on the number of independent variables, the regression equation can be either one-variable or multiple regression. If the regression function is linear, the variables are said to be linearly correlated. Univariate linear regression analysis, includes only one independent variable and one dependent variable [11]. The relationship between the two can be approximated by a straight line, with the independent variable denoted by x . The dependent variable (predictor variable) is denoted by y . The expression takes the form $y = w' + e$, with e being the error obeying a normal distribution with a mean of 0. The correlation coefficient is a statistical indicator first devised by statistician Carl Pearson. It is a quantity that studies the degree of linear correlation between variables, generally denoted by the letter r . There are various ways to define the correlation coefficient due to the different objects of study, and the more commonly used one is Pearson's correlation coefficient.

$$r(X, Y) = \frac{COV(X, Y)}{\sqrt{Var[X]Var[Y]}} \quad (1)$$

$Cov(X, Y)$ is the covariance of X and Y , $Var[X]$ is the variance of X , $Var[Y]$ is the variance of Y . Defining the

$$\rho_{XY} = r(X, Y) \quad (2)$$

One notices that the absolute value is smaller than 1. When $|\rho_{XY}| = 1$, there is a set of constant groups $[a, b]$ satisfying $P\{Y = a + bX\} = 1$. The correlation coefficient quantitatively portrays the degree of correlation between X and Y , i.e., the more significant the $|\rho_{XY}|$, the greater the correlation; $|\rho_{XY}| = 0$ corresponds to the lowest degree of correlation. X and Y are perfectly correlated when there is a linear relationship with probability one and is then a quantity that characterizes the closeness of the linear relationship between X and Y . When $|\rho_{XY}|$ is large, X and Y are well correlated; when $|\rho_{XY}|$ is small, X and Y are poorly correlated; when X and Y are not correlated, the assumption indicates no linear relationship between X and Y . However, it does not exclude those other relationships may exist between X and Y .

For analysis software, SPSS is an analysis tool in this research, which is also named for Statistical Product Service Solutions, a statistical product and service solution software. The essential functions of SPSS include data management, output management, graphical analysis, statistical analysis, and

more. In this study, this software output the general linear model, ANOVA test, and regression analysis.

Table I. Model summary of R evaluation, F change, and df.

R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
				R Square Change	F Change	df1	df2	Sig. F Change
.985 ^a	.971	.968	1.57995526	.971	397.179	3	36	.000

3. Results & Discussion

As shown in Table. II, the Correlation coefficient table, statistical table of regression results, and predicted value - actual value are presented. Correlation is significant at the 0.01 level (2-tailed). According to the table above, there is a significant Pearson correlation between all tertiary courses and the bachelor's, master's, and doctorate courses in China, irrespective of the student's country of origin. It indicates that students from all works of life and levels find China an exciting country to learn, which can be attributed to stable economic growth, peace, advanced information, and Communication Technology.

Table II. Correlation coefficients.

	All tertiary	Bachelor	Masters	Doctor
All tertiary	1	0.929	0.969	0.846
Bachelor	0.929	1	0.882	0.743
Masters	0.969	0.882	1	0.817
Doctor	0.846	0.743	0.817	1

As shown in Table. III, the dependent Variable is international students undertaking tertiary courses in China. The Predictors are Constant, showing international students undertaking doctoral courses in China, International students undertaking bachelor's courses in China, and international students undertaking master's courses in China.

Table III. Statistical table of regression results.

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	2974.381	3	991.460	397.179	.000
Residual	89.865	36	2.496		
Total	3064.247	39			

As shown in Fig. 2, the regression shows a positive relationship between international students studying tertiary courses and those undertaking master's degrees. On this basis, this paper concludes that China has a higher affinity for attracting students from OECD countries for courses of tertiary, bachelor, masters, and doctoral studies.

As shown in Fig. 3, There is a weak positive relationship between international students in China and Doctoral studies. It implies that the number of cross-border students is fewer at the doctoral level. Dynamic analysis reasonably infers that Chinese tertiary education's knowledge structure, course design, economic stability, and technological advancement are at least 74.3% attractive to foreign international students studying in China. Therefore, an attractive marketing strategy needs to be conducted to increase the proportion of intended foreign international students' input.

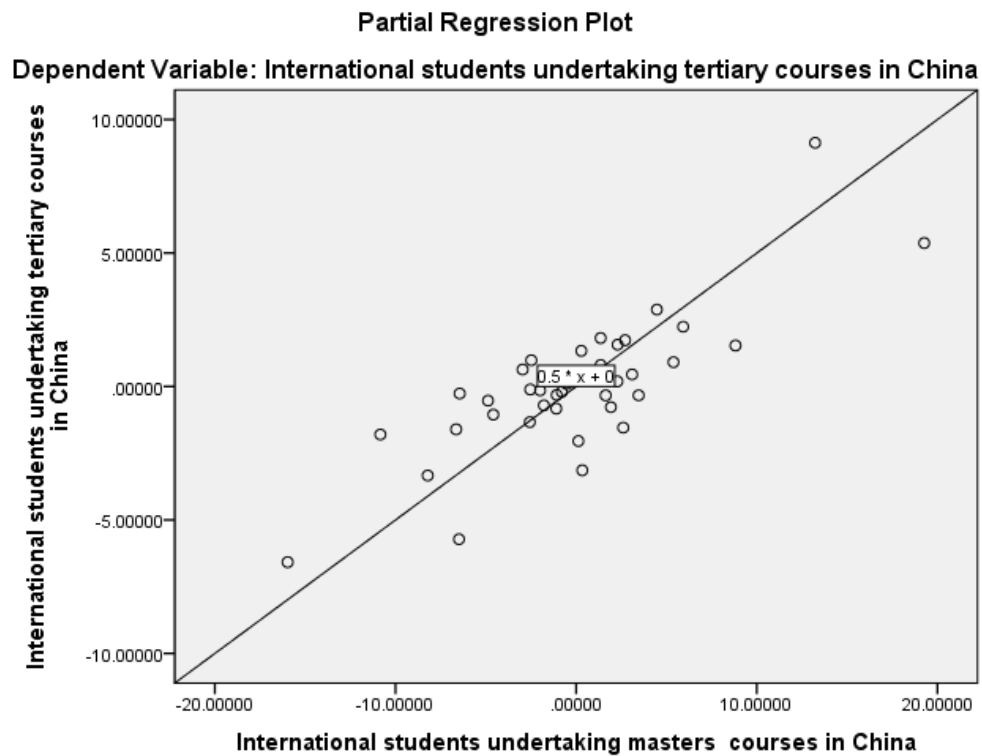


Fig 2. Partial regression plot of international students, undertaking masters' courses in China.

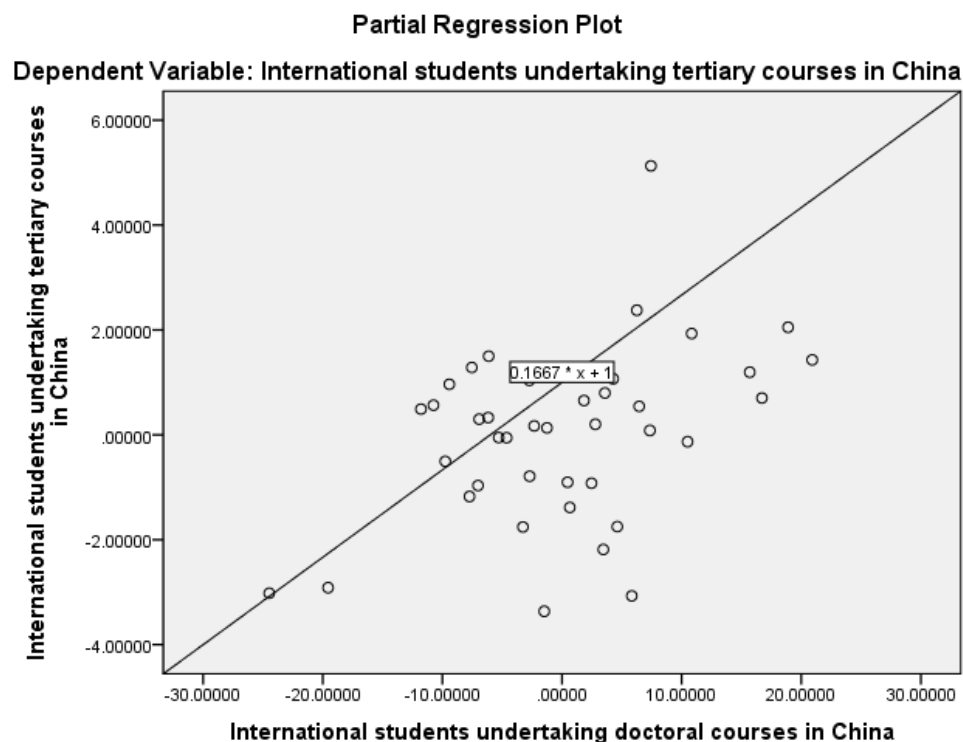


Fig 3. Partial regression plot of international students, undertaking doctoral courses in China.

In static analysis, according to the Planned behavior theory (TPB theory), the consumer's consumer behavior is immediately decided by intention. Intention tributes mainly include attitude towards the behavior, subjective norms, and behavior control. Furtherly derivate, the motivators are 1. The belief that the behavior led to specific outcomes 2. Evaluation of outcomes 3. Beliefs of important referents 4. Motivation to comply with referents. To make Chinese tertiary education more attractive to foreign

international students, researchers need to make a reasonable plan to address these four issues. In the belief of behavior, outcomes evaluation is the most important motivator. Institutional branding marketing is of great importance to optimize the consumers' intended outcomes.

3.1 Established Brand Layout

Develop a general understanding of students. Analyze the students' proportion, Define the value of branding towards the most likely intended students who will provide the most advancement towards school, and embody value characteristics into the band.

3.2 Internalize brand traits into inner employees

After setting up the brand's general layout, Internalize the brand personality into specific employees. In the Recruitment and training process, insert the brand value into every employee, and make the employee correspond to the brand.

3.3 Penetration campus culture for students

Although educational institutions encourage self-learning and group learning, which segment students into smaller groups, culture penetration is also essential to condensate students.

3.4 Alumni management

Students' statues are the best communication media on campus. Harvard University began to carry out strict alumni management in the early days of its establishment, while India and China, and other developing countries began to cultivate alumni relations in the early 20th century consciously. Alumni relationship cultivation can make effective publicity for the school and enable the school to establish a parallel relationship with the business built by alumni. With the parallel relationship between school and business, it is easier for schools to raise research funds through school-enterprise donations and partnerships with businesses. Many research funds can be invested in excellent students to achieve research results. In the long run, the inflow of funds, the investment of funds, the output of results, the increase of soft and hard power of the university, and the university's world ranking will be improved. These six links will be closely combined to form a benign development loop.

According to the Psychoanalytic Theory Desires & Emotions, the behavior motivator is beyond income, age, and country, which is hugely affected by emotions, aspirations, fears, feelings and desires. Therefore, educational institutions' marketers could establish irrespective marketing campion to stimulate the potential effects. To stimulate the feelings, the educational institution could implement the following method.

3.5 Use school heritage to embellish and promote the educational institution's brand

The heritage of a university establishes a distinctive first impression toward intended students [12]. The profound school should take advantage of their school heritage to optimize the school's sedate feeling, which implies better-intended outcomes in international students' judgment. According to Vilniaus university's marketing strategy, the new educational institution could also use traditional attributes, like museums, ceremonies, and old elements of the insignia, to increase profound culture's implication.

3.6 Shape educational institutional identity

Shaping educational institutional identity help shape foreign international students' recognition of cultural identity. Christian Higher Education states that faith-based branding marketing significantly develops international foreign admissions [13]. This rule is not only suited for faith-based universities but also helps non-faith-based universities to grow. The non-faith-based universities could develop a cultural-friendly environment for foreign international students.

In static analysis, based on all assumptions in research, any foreign countries' international Educated students' marketing indicators changes will impact their Chinese study consumer behavior.

The international students' consumer preference for foreign countries is still subject to change beyond the assumed educational institution context. Students' marketing assumptions non-concerned in this research may still lead to significantly different consumer performance. Therefore, this integrated consumer motivation model still needs to be considered derived from weighted assumptions of proprietary consumer preferences.

Secondly, weighted sorting is needed in an integrated model because different factor has a different degree of influence. If the researchers combine all foreign international student's preference attributes into the same model, the preference exhibition will be intuitive but dilutive because some indicators have Duality. Specifically, a newly Chinese educational institution charges a cheaper tuition fee but incomplete knowledge frameworks; the cheaper tuition fee is a positive attribute that attracts foreign international students. However, an incomplete knowledge framework is a Negative momentum, which generates a downward trend but is diluted by the non-weighting integrated model. Therefore, researchers need to assign a higher weight to more Positive preference indicators towards Chinese educational institutions to mitigate unintended results and purify intended preference results. Factor analysis is recommended.

Thirdly, in dynamic analysis, Pearson correlation interpretation only demonstrated the linear correlations. However, Fishbein & Ajzen have proposed that consumer behavior formation experienced complicated psychological processes. Therefore, the correlation coefficient table cannot reflect a concrete relationship between tertiary courses and foreign international students' bachelor's, master's, and doctorate course preferences toward China entirely; Further relationship analysis still needs to be conducted.

4. Conclusion

In summary, this paper explores and investigates the foreign international student's attitudes toward Chinese education mainly based on statistical prediction and marketing strategy to. Statistical predictions are concerned with Correlation coefficient, R square, ANOVA test, and Partial regression plot. The Independent variable is education, while the dependent variables include technological advancements, consumer behavior, institution brand marketing, The static analysis includes Planned behavior theory (TPB theory) and Psychoanalytic Theory Desires & Emotions. In corresponding behavior, educational institutions marketers should establish the Institutional brand marketing and promotional campaign to increase intention and substantive emotional transformation. According to the analysis, this paper provides a good vision for foreign international students on how to react to Chinese educational institutions' future advancement. In the future, weighted sorting and further assumption should be made for static analysis; more complex correlation interpretation should be combined into the research for dynamic analysis. Overall, these results offer a guideline for Chinese educational institutions to develop their marketing strategy.

References

- [1] G. C. Textor. "Leading destinations for Chinese students studying abroad 2015 and 2021." Statista. Last modified July 21, 2021. Accessed May 3, 2022. Available at: <https://www.statista.com/statistics/1108708/china-students-preferred-destinations-for-study-abroad/>.
- [2] B. B. Sousa, and F. C. Magalhães. "An Approach on Attachment in Public Marketing and Higher Education Management Contexts." Higher Education and the Evolution of Management, Applied Sciences, and Engineering Curricula, vol. 1 2019, pp. 151–171.
- [3] W. Hu, and Y. Shi. "Prediction of Online Consumers' Buying Behavior Based on LSTM-RF Model." 2020 5th International Conference on Communication, Image and Signal Processing (CCISP), 2020.
- [4] G. Yang, and P. Zheng. "A Prediction Method of Consumer Behavior Transformation in K12 Educational Institutions." Frontiers in Educational Research. Francis Academic Press, n.d. Accessed May 3, 2022.
- [5] A. Ang, "Multifactor Strategies." BlackRock. Accessed May 3, 2022. Available at: <https://www.blackrock.com/institutions/en-axj/insights/factor-perspectives/multi-factor-strategies>.

- [6] N. Stebliuk, and N. Kuzmenko. "Research of Consumer Demand in the Market of Educational Services of Dnipropetrovsk Region." *Economies' Horizons*, vol. 3(14), 2021, pp. 64–71.
- [7] I. Ajzen, "Consumer Attitudes and Behavior." *Handbook of Consumer Psychology*, 2012.
- [8] H. Baumgartner, and J. E. M. Steenkamp. "Exploratory Consumer Buying Behavior: Conceptualization and Measurement." *International Journal of Research in Marketing* vol. 13, no. 2 1996, pp. 121–137.
- [9] OECD/UIS/Eurostat, Table B6.1. See Source section for more information and Annex 3 for notes 2021. Available at: https://www.oecd.org/education/education-at-a-glance/EAG2021_Annex3_ChapterB.pdf.
- [10] B. O. Рахманов, "Methodical Bases of Educational Information Environment Formation in the Technical Higher Educational Institutions." *Proceedings of the National Aviation University. Series: Pedagogy, Psychology*, vol. 7, 2015.
- [11] N. T. Thomas, "A LSTM Based Tool for Consumer Complaint Classification." 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2018.
- [12] N. Bulotaite, "University Heritage—an Institutional Tool for Branding and Marketing." *Higher Education in Europe* vol. 28, no. 4, 2003, pp. 449–454.
- [13] D. Tolbert, "An exploration of the use of branding to shape institutional image in the marketing activities of faith-based higher education institutions." *Christian Higher Education* vol. 13.4, 2014, pp. 233-249.