Salud, Ciencia y Tecnología - Serie de Conferencias. 2025; 4:1787
doi: 10.56294/sctconf20251787

REVIEW

Factors Influencing The Drop-Out Rate Of International Students In South Korea: Application Of Big Data, Social Network, And Machine Learning Analyses

Factores que influyen en la tasa de abandono escolar de estudiantes internacionales en Corea del Sur: Aplicación de análisis de big data, redes sociales y aprendizaje automático

Young-Chool Choi[1] ✉, Yun Na Jeong[2]

[1]Dept. of Public Administration, Chungbuk National University, Korea.
[2]UNIST, Korea.

Cite as: Young-Chool C, Na Jeong Y. Factors Influencing The Drop-Out Rate Of International Students In South Korea: Application Of Big Data, Social Network, And Machine Learning Analyses. Salud, Ciencia y Tecnología - Serie de Conferencias. 2025; 4:1787. https://doi.org/10.56294/sctconf20251787

Submitted: 21-11-2024 **Revised:** 12-01-2025 **Accepted:** 28-02-2025 **Published:** 01-03-2025

Editor: Dr. William Castillo González

Corresponding author: Young-Chool Choi ✉

ABSTRACT

This study aims to identify the key factors influencing the attrition rate of international students in South Korea by utilizing big data analysis, machine learning, and social network analysis. By analyzing 643 academic papers related to international students, this research seeks to propose measures to improve their retention rates. Furthermore, machine learning techniques are employed to pinpoint the most significant variables affecting student attrition. The analysis covers 221 four-year universities in Korea, focusing on variables such as the type of institution, student capacity, and geographical location. Results show that these universities can be categorized into seven nodes, with the admission competition rate of new students being the most influential factor; higher competition correlates with lower attrition rates. This study enhances objectivity and validity by employing advanced statistical techniques not used in previous research, providing scientific evidence to support policy improvements.

Keywords: Big Data Analysis; Machine Learning; Social Network Analysis; International Students; Attrition Rate.

RESUMEN

Este estudio busca identificar los factores clave que influyen en la tasa de deserción estudiantil internacional en Corea del Sur mediante el análisis de big data, aprendizaje automático y análisis de redes sociales. Mediante el análisis de 643 artículos académicos relacionados con estudiantes internacionales, esta investigación busca proponer medidas para mejorar sus tasas de retención. Además, se emplean técnicas de aprendizaje automático para identificar las variables más significativas que afectan la deserción estudiantil. El análisis abarca 221 universidades de cuatro años en Corea, centrándose en variables como el tipo de institución, la capacidad estudiantil y la ubicación geográfica. Los resultados muestran que estas universidades pueden clasificarse en siete nodos, siendo la tasa de competencia para la admisión de nuevos estudiantes el factor más influyente; una mayor competencia se correlaciona con menores tasas de deserción. Este estudio mejora la objetividad y la validez mediante el empleo de técnicas estadísticas avanzadas no utilizadas en investigaciones previas, lo que proporciona evidencia científica para respaldar la mejora de las políticas.

Palabras clave: Análisis de Big Data; Aprendizaje Automático; Análisis de Redes Sociales; Estudiantes Internacionales; Tasa de Deserción.

INTRODUCTION

Amidst global demographic changes and labor market imbalances, South Korea faces continuous population decline and aging issues. Attracting international students has become a crucial strategy for maintaining national competitiveness and revitalizing the labor market. Particularly as the school-age population decreases, the operation of universities is becoming challenging, with some regional institutions facing existential threats due to recruitment difficulties. Actively recruiting international students and enhancing their retention is vital for addressing these issues.[1]

The attrition rate of international students in Korean universities remains high, presenting an urgent challenge that must be addressed at both national and regional levels. High attrition rates can lead to academic failures for the students, diminish universities' global reputations, exacerbate regional economic and depopulation issues, and signify a failure to secure global talent.[2] Especially for regional universities, maintaining international student enrollment is crucial for preventing regional depopulation and contributing to economic revitalization. Hence, empirical and multifaceted analyses are necessary to reduce international student attrition rates.

Existing studies on the attrition of international students have predominantly relied on qualitative methods like surveys and interviews or simple statistical analyses, which are limited in analyzing the comprehensive factors affecting students' continuation of studies.

With advancements in big data analytics and artificial intelligence (AI), there is a need for new methodological approaches that can quantitatively analyze patterns of student attrition.[3,4,5] Sophisticated analytical methods such as Big Data Analysis, Social Network Analysis, and Machine Learning can intricately explore the diverse factors causing attrition. For example, machine learning can effectively predict attrition using demographic and academic data of students, while social network analysis can elucidate the impact of students' social relationships and network structures on attrition.

However, research utilizing these advanced analytical methods in a combined approach is scarce, and previous studies have been limited to isolated analyses using individual techniques, which have not thoroughly interpreted the complex relationships among factors influencing attrition. Therefore, this study aims to deeply understand the issue of international student attrition and derive practical improvement measures by integrating various analytical methodologies.

This study aims to identify the main factors influencing the attrition rates of international students studying in Korea and to propose effective improvement measures based on these factors. By employing big data analysis, social network analysis, and machine learning in a combined approach, this research predicts the attrition rates of international students, explores influencing factors, and seeks measures to reduce attrition rates at the university and policy levels.

Specifically, the main objectives of this study are to:

1) Identify factors influencing international student attrition using big data and social network analysis methods.

2) Analyze the relationships among key influencing variables of student attrition using social network analysis.

3) Develop a predictive model for student attrition using machine learning.

4) Based on the analysis results, propose practical policy and institutional improvements at university and governmental levels to reduce international student attrition.

This research is expected to contribute academically by introducing new analytical methodologies to attrition studies and practically by providing empirical evidence needed for universities and policy-making bodies to develop more effective international student support systems.

Theoretical Background and Review of Previous Studies

Importance of Managing International Students

Globally, universities are enhancing their international student management systems to strengthen global competitiveness and attract international talent. Managing international students goes beyond providing education; it plays a crucial role in the sustainable development of universities, local communities, and the nation as a whole.

International students often find employment or start businesses in the host country after their studies, potentially solving labor shortages and contributing to the national economy. Particularly in societies with low birth rates and aging populations, like Korea, attracting highly skilled personnel is essential, and managing international students is a critical process in this talent inflow.[6] For example, in the United States, international students contribute to the local workforce supply, especially in STEM fields.[6]

From the perspective of enhancing universities' global competitiveness, world universities consider the ratio of international students as one of the key metrics in global rankings such as the QS World University Rankings and THE (Times Higher Education) rankings. Therefore, actively recruiting and effectively managing international students plays a significant role in enhancing the brand value and rankings of universities.[7]

Particularly for regional universities, international students play a crucial role in revitalizing local economies

and addressing depopulation. Many students find employment and settle in the region after their studies, having a long-term positive impact on the community.[8] Countries like Japan, Germany, and Canada consider this and actively support the settlement of international students.

Lastly, international students often return to their home countries after their studies, where they facilitate diplomatic and economic exchanges between the countries. The Korean government is also promoting strategies to attract students through the 'Study Korea' policy linked with Korean Wave and culture, while countries like the USA and UK also focus on cultural diplomacy and building global networks through international students.[9]

Comparative Analysis of International Student Management Policies in Major Countries

As the internationalization of higher education accelerates globally, major countries are implementing various policies to attract international students. Each country's policies are closely linked to its educational environment, economic conditions, labor market demands, and immigration policies, and their strategies for attracting students also show distinctive features.

The USA attracts the largest number of international students worldwide, with over one million students enrolled in American universities as of 2023. US universities rank high in global university evaluations like the QS Rankings and THE World University Rankings, especially showing strong competitiveness in STEM fields.[6] However, tuition fees are relatively high, and international students have the opportunity to work in the USA for a certain period after graduation through the Optional Practical Training (OPT) program.[10]

The UK, with its long academic tradition and globally recognized university system, especially with prestigious universities like Oxford and Cambridge, has seen a surge in non-EU student ratios since Brexit in 2019. The UK government introduced the Graduate Route, allowing graduates to stay and work for up to two years (three years for PhD holders) after graduation.[11] Major countries of origin for students in the UK include China, India, Nigeria, and Malaysia, with a relatively high proportion of graduate and doctoral students.

Germany offers an attractive educational environment for international students, with most public university tuition fees being free or very low. Additionally, Germany has a strong vocational training system and industry-linked programs, offering many opportunities for students to find employment within the country after their studies. The German government actively supports the stay and employment of international students, allowing them to stay for up to 18 months after graduation to prepare for employment.[12] Major source countries include India, China, Turkey, and Russia, with a high proportion of students in engineering and STEM fields.

Canada is one of the countries actively promoting student-friendly policies, encouraging the "Study-to-Immigration" pathway that allows international students to obtain permanent residency after their studies. Students can apply for the Post-Graduation Work Permit (PGWP), which allows them to work for up to three years after graduation, linking it to Canada's immigration programs (PNP, Express Entry) for long-term residency. Major source countries for Canadian international students include India, China, the Philippines, and Nigeria, and due to relatively low tuition fees and high quality of life, the number of students has been increasing. Canada's policies focus on supporting graduates to contribute to Canadian society after their studies.[13]

Australia, with a high proportion of students from Asia, offers programs that combine English language studies and university education. China, India, Vietnam, and Nepal are among the major countries of origin, and the Australian government is expanding support for employment and residency for international students. After meeting certain conditions, graduates can work through the Temporary Skill Shortage (TSS) visa, with some fields linked to pathways for obtaining permanent residency. Although Australia maintains relatively open policies for international students, the increasing cost of tuition and living expenses is prompting a reassessment of its recruitment strategies.[14]

Japan is actively expanding its policies to attract international students, with the government's goal of enrolling 300 000 students by implementing the 'Global 30' project. Particularly, students from Korea, China, Vietnam, and Nepal have high proportions, not only in universities but also in vocational training institutes and Japanese language education institutions. The Japanese government supports the employment and residency of international students, especially promoting employment in IT and engineering fields after graduation. However, Japanese language proficiency plays a critical role in employment opportunities, which can be a burden for students.[10]

France, with its low tuition fees and rich French education programs, is an attractive option for international students. Major source countries include Morocco, China, Algeria, and Senegal, particularly with a high proportion of students in fields like culture, arts, design, and social sciences. The French government is expanding support for students' stay, allowing them to stay for up to two years after graduation. However, non-EU students face more procedures for extending their stay and obtaining work visas, which can be a disadvantage.

3. Comparative Analysis of International Student Policies by Country and Korea's Policy

Major countries worldwide attract international students in different ways and operate policies supporting their academic continuity and settlement after graduation. The USA and the UK, based on their global university

https://doi.org/10.56294/sctconf20251787

competitiveness, continue to attract many students despite high tuition fees, while Germany and France actively recruit students through low tuition and public support. Canada and Australia induce long-term stays of graduates through policies linked to employment and immigration, and Japan operates student policies linked to Japanese education, focusing on students from the Asian region.

Korea aims to attract 300 000 international students by 2027 through the "Study Korea 300K" project, with the internationalization of higher education and enhancing the competitiveness of regional universities as major tasks. Korean policies and current practices in managing international students compared to major countries show the following characteristics:

(1) Policies for Attracting International Students and Academic Support

The Korean government operates scholarship programs like the Global Korea Scholarship (GKS) to attract international students, and universities also offer various scholarships for international students.

Korean universities run Korean Language Programs for international students, and some expand English-only lectures and operate International Colleges to create a student-friendly educational environment. However, students with insufficient Korean language skills often struggle to maintain their academic studies, and the level of support for international students varies among universities, with some lacking systematic management.

(2) Policies Supporting Residency and Employment for International Students

The Korean government provides employment visas such as D-10 (Job Seeking Visa) and E-7 (Specific Activities Visa) to enable students to settle in Korea after graduation. Recently, policies have been revised to recognize graduates meeting certain criteria as "excellent foreign talents," allowing for long-term residency and permanent residency acquisition.

However, Korean companies are somewhat reluctant to hire international students, and the opportunities for employment in Korea after graduation are relatively limited.

(3) Support for International Students' Daily Life and Community Integration

Korean universities offer dormitories, health insurance enrollment support, and life counseling services for international students, but programs for systematic integration into the community are relatively lacking. There is a growing disparity in the proportion of international students between metropolitan and regional universities, and although regional universities are expanding benefits such as tuition reduction and dormitory provision to attract students, the retention rate of international students remains low. Some local governments operate programs to support the settlement of international students to address regional economic and population decline issues, but their effectiveness has not yet been fully proven.

Review of Previous Studies and the Uniqueness of This Research

Research on the management policies of international students has primarily focused on topics such as recruitment strategies, factors influencing continuation and dropout, and post-graduation residency and employment policies. Previous studies have discussed the following key areas:

Firstly, studies on the recruitment and educational environment improvement for international students have analyzed the strategic and policy differences in attracting students internationally. Altbach & Knight[15] analyzed major trends in international higher education, emphasizing the intensifying competition among countries to attract students. According to their research, while countries like the USA and UK attract students based on their global university competitiveness, Germany and France draw students by reducing tuition fees. The OECD[16] report suggests that student recruitment policies are linked to national economies and labor markets and are deeply related to long-term immigration policies.

Secondly, studies on the factors contributing to student dropout and academic persistence have identified language barriers, academic stress, cultural adaptation, and economic factors as major contributors. Wilkins et al.[17] emphasized the need for universities to strengthen academic and social support to reduce dropout rates, particularly support during the initial adaptation phase. Andrade[18] found that students often drop out due to linguistic difficulties and academic pressures, noting that academic persistence increases when support from faculty and administrative bodies is effectively provided.

Thirdly, research on post-graduation residency and employment support policies has addressed the pathways and settlement possibilities for students after graduation. Hawthorne[19] compared student policies in Canada, Australia, and New Zealand, analyzing how providing employment opportunities to students positively impacts the national economy and labor market. Gribble[20] pointed out that while policies allowing post-graduation stay enhance the competitiveness in attracting students, the willingness of companies to hire international students is a crucial variable for policy success.

Previous studies primarily analyzed the policies of international students in single countries or evaluated the impact of specific policies on students' academic persistence and employment based on qualitative analysis or

surveys. However, this research is distinctive in the following ways:

This study, first, combines big data analysis, social network analysis (SNA), and machine learning analysis[21,22,23,24] to explore various factors affecting the dropout rates of international students from multiple angles. While previous studies relied mainly on surveys and statistical analysis, this research employs a quantitative approach using large-scale data for more objective and in-depth analysis.

Secondly, it proposes a model for predicting the dropout rates of international students and deriving policy implications. Specifically, it develops a predictive model using machine learning to provide practical management strategies that can be utilized by universities and policymakers.

Therefore, this research aims to address the limitations of previous studies and contribute to improving the management policies of international students by applying a new research methodology that emphasizes a quantitative approach.

This research aims to identify the key factors influencing the dropout of international students in Korea by using big data analysis, social network analysis, and machine learning analysis on academic papers related to 'foreign students' published in Korea.

METHOD

The overall design of the study is as follows:

Research Subjects and Analysis Data

This study analyzes 643 academic papers related to 'foreign students' registered with the Korea Citation Index (KCI) from January 1, 2010, to December 31, 2024. The criteria for selecting the papers are as follows:

- Papers with the term 'foreign students' in the title are selected.
- Paper data is collected based on the KCI database.
- Duplicate papers and those not directly related to the research scope are excluded.

This research uses the titles, abstracts, and keywords data of these papers for big data analysis and social network analysis, applying the major concepts derived to machine learning analysis for an in-depth analysis of dropout factors.

Research Method and Analysis Procedure

This study uses Netminer 4.5 and IBM Modeler 18 programs to apply big data analysis, social network analysis (SNA), and machine learning analysis (decision tree analysis). The analysis procedures are as follows:

(1) Phase 1: Big data analysis to analyze trends in studies related to international students.

The titles, abstracts, and keywords from the 643 papers are extracted for text data cleaning and preprocessing. Key word frequency and co-occurrence analysis are performed to derive important concepts in the research of international students, and text mining techniques are used to identify major research topics and trends.

(2) Phase 2: Using SNA to derive key networks related to dropout.

From the results of the big data analysis, an ego-network is created focusing on keywords related to 'Attrition, Dropout.' PFnet (Pathfinder Network) analysis is conducted using Netminer 4.5. Centrality analysis and connection structure analysis are performed to derive concepts closely related to the dropout of international students, and community detection techniques are applied to derive major concept groups (communities) related to dropout.

(3) Phase 3: Hypothesis setting on key influencing factors of dropout.

Based on the network analysis results, variables likely to influence the dropout of international students are derived. Hypotheses are set using the results of big data and social network analysis on Korean academic papers, assuming factors that could influence dropout.

Hypothesis 1: National universities will have a lower dropout rate of international students compared to private universities.

Hypothesis 2: Universities in the capital region will have a lower dropout rate of international students compared to those outside the capital region.

Hypothesis 3: Universities with a larger admission quota will have a lower dropout rate of international students.

Hypothesis 4: Universities with a higher competitive rate of new student admissions will have a lower dropout rate of international students.

Hypothesis 5: Universities with a higher number of international students will have a lower dropout rate.

Hypothesis 6: Universities with higher degrees of exchange with foreign universities will have a lower dropout rate of international students.

Hypothesis 7: Universities with stronger capacities for community service will have a lower dropout rate of international students.

°(4) Phase 4: Verification of key variables using machine learning analysis (Decision Tree Analysis)

https://doi.org/10.56294/sctconf20251787

Machine learning analysis is conducted based on the variables set in the hypotheses. Decision Tree Analysis is performed using IBM Modeler 18.

(5) Variables: the dependent variable is the dropout rate of international students at the universities, and the independent variables are as follows:

- Type of establishment: whether the university is founded by the government (public=1) or by private entities (private=2).
- Region: whether the university is located in the capital region (capital=1) or outside the capital region (non-capital=2).
- Admission quota: the size of the school's admission quota.
- Admission competition rate: the competitive rate of new student admissions.
- Number of international students: the number of international students.
- Degree of exchange with foreign universities: the number of students sent to foreign universities.
- University's capacity for community service: the ratio of students completing community service to the total number of students.

RESULTS

Results of Word Cloud, PFnet, and Clustering Analysis Firstly, the results of the word cloud analysis using the words derived from the 643 academic papers are as follows.

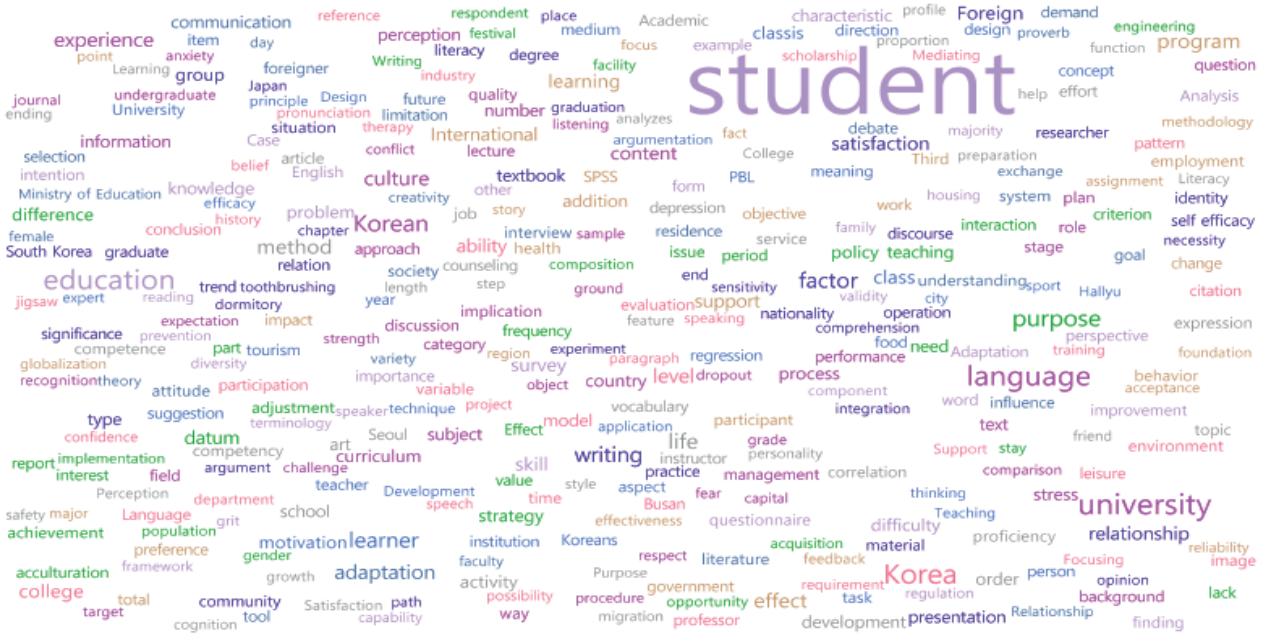

Figure 1. Word Cloud Analysis. Note: Includes only the top 500 words by frequency.

The number of academic papers that include 'foreign students' in their titles amounts to 643, authored by a total of 742 individuals. The data included in this analysis consists of the English abstracts of academic papers published from 2010, when policies regarding foreign students began to be formally established in Korea, until December 31; 2024. The analysis includes 4 437 words. The frequency of these words, sorted in order, is shown in the table (figure 2).

To simplify the network composed of these words, an ego-network including 'dropout' was constructed, and PFnet analysis was conducted using this ego network to reorganize the complex network structure into a simpler form. The streamlined network structure obtained through PFnet analysis is shown in figure 3.

Using the network derived from the PFnet analysis, the community structure related to the dropout of international students was further extracted by clustering.

2. Identification of Key Variables

The two network analysis results presented above (PFnet analysis and clustering results) visually illustrate the relationships among various factors influencing the dropout rate of international students. Below, each variable is explained in connection with the key concepts and relationships derived from the networks.

First, University Type (Public vs. Private)

https://doi.org/10.56294/sctconf20251787

		1	2	3	4	5
		Part of Speech(POS)	Frequency	Word length	Name Type	Author Keyword
1	student	"Common Noun"	3,767.0	7.0	"-"	"False"
2	university	"Common Noun"	758.0	10.0	"-"	"False"
3	language	"Common Noun"	695.0	8.0	"-"	"False"
4	education	"Common Noun"	687.0	9.0	"-"	"False"
5	Korea	"Proper Noun"	634.0	5.0	raphical Name"	"False"
6	purpose	"Common Noun"	433.0	7.0	"-"	"False"
7	Korean	"Proper Noun"	421.0	6.0	nization Name"	"False"
8	factor	"Common Noun"	372.0	6.0	"-"	"False"
9	learner	"Common Noun"	359.0	7.0	"-"	"False"
10	life	"Common Noun"	356.0	4.0	"-"	"False"
11	writing	"Common Noun"	338.0	7.0	"-"	"False"
12	program	"Common Noun"	317.0	7.0	"-"	"False"
13	method	"Common Noun"	313.0	6.0	"-"	"False"
14	effect	"Common Noun"	305.0	6.0	"-"	"False"
15	culture	"Common Noun"	301.0	7.0	"-"	"False"
16	adaptation	"Common Noun"	282.0	10.0	"-"	"False"
17	experience	"Common Noun"	276.0	10.0	"-"	"False"
18	level	"Common Noun"	265.0	5.0	"-"	"False"
19	class	"Common Noun"	256.0	5.0	"-"	"False"
20	college	"Common Noun"	255.0	7.0	"-"	"False"
21	course	"Common Noun"	251.0	6.0	"-"	"False"
22	learning	"Common Noun"	242.0	8.0	"-"	"False"
23	skill	"Common Noun"	240.0	5.0	"-"	"False"
24	Foreign	"Proper Noun"	230.0	7.0	nization Name"	"False"
25	satisfaction	"Common Noun"	226.0	12.0	"-"	"False"
26	content	"Common Noun"	219.0	7.0	"-"	"False"
27	survey	"Common Noun"	217.0	6.0	"-"	"False"
28	relationship	"Common Noun"	206.0	12.0	"-"	"False"
29	International	"Proper Noun"	195.0	13.0	nization Name"	"False"
30	datum	"Common Noun"	194.0	5.0	"-"	"False"

Figure 2. Frequency of Appearing Words

It is essential to verify whether keywords such as “public,” “government,” “private,” and “finance” appear in the network analysis. If concepts like “support,” “program,” “scholarship,” and “stability” are closely associated with “public university,” it suggests that public universities may receive more stable financial support and thus have lower dropout rates. Conversely, if “private” is linked to “cost” and “financial difficulty,” it indicates that the high tuition burden of private universities may contribute to international student dropouts. In this network analysis, public universities likely have more government support, lower financial burdens, and more support programs, which could lead to lower dropout rates.

Second, University Location (Capital Region vs. Non-Capital Region)

It is necessary to check whether words such as “Seoul,” “region,” “environment,” “opportunity,” and “support” appear in the analysis. If “Seoul” is linked to “opportunity,” “exchange,” and “adaptation,” it implies that universities in the capital region may offer a better environment for international students. Conversely, if “region” is connected to “difficulty,” “lack,” and “isolation,” it suggests that regional universities may face support deficiencies that increase dropout rates. The associated network elements indicate that capital region universities provide better opportunities and networks, helping international students adapt more easily and reducing dropout rates.

https://doi.org/10.56294/sctconf20251787

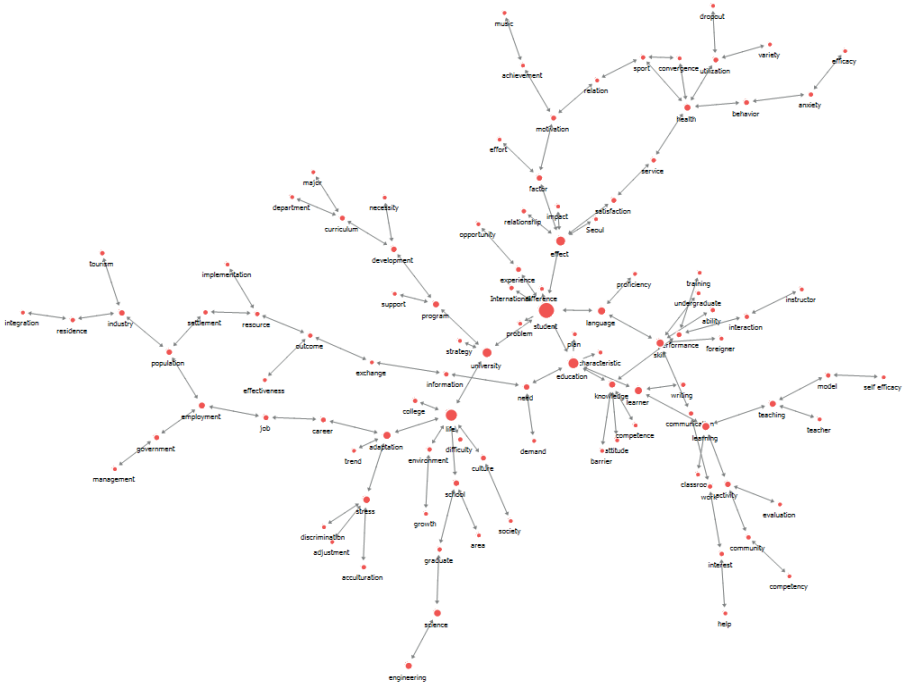

Figure 3. Results of PFnet Analysis

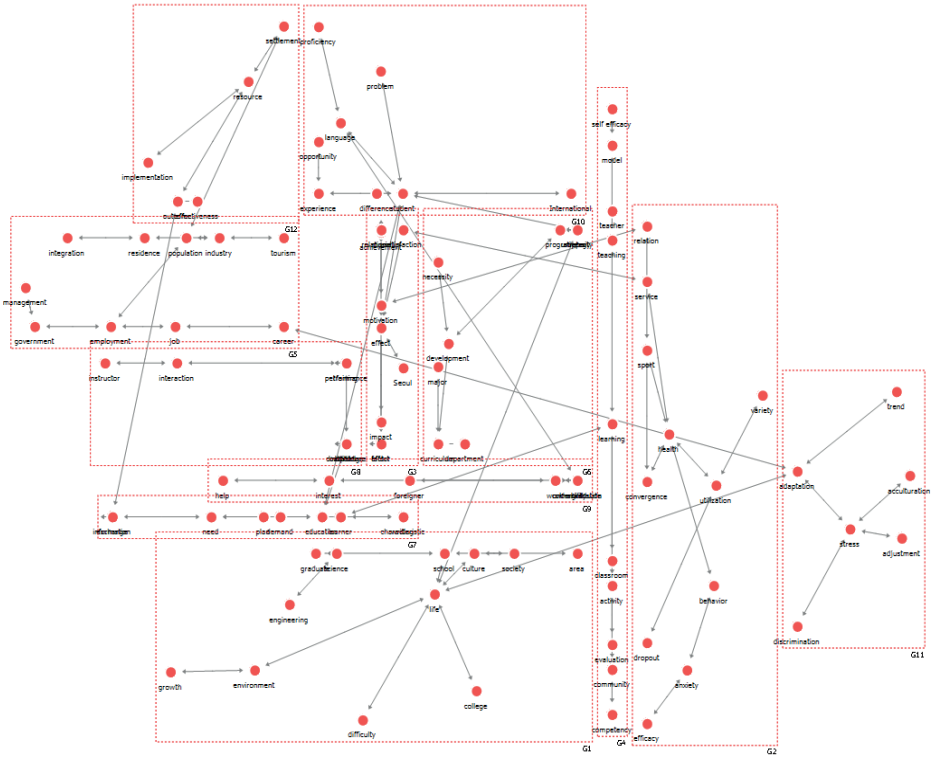

Figure 4. Community Structure Derived from Clustering

Third, University Admission Quota (University Size)

It is necessary to check whether terms like “admission,” “quota,” “large scale,” “small scale,” and “resource” appear in the network. If “quota” is linked to “support,” “program,” and “efficiency,” it implies that universities with larger admission quotas may have more structured support systems for international students. However, if “large scale” is connected to “problem” and “inefficiency,” it suggests that simply having a large admission quota does not necessarily lower dropout rates. Given the associated network elements,

https://doi.org/10.56294/sctconf20251787

universities with larger quotas are likely to have better support systems for international students.

Fourth, Admission Competition Rate

It is important to check whether words such as “competition,” “admission,” “level,” and “excellent student” are linked in the network. If highly competitive universities are associated with “excellent student” and show strong connections with concepts like “success” and “adaptation,” it suggests that universities with higher admission competition rates may have lower dropout rates. Based on the related network elements, universities with higher competition rates likely attract higher-achieving students, which may contribute to lower dropout rates among international students.

Fifth, Number of International Students

It is necessary to verify whether words such as “foreign student,” “exchange,” “community,” and “support” are closely connected in the network. If “foreign student” is linked to “community,” “adaptation,” and “support,” it suggests that universities with a large number of international students have well-developed networks that help reduce dropout rates. Based on the network elements, universities with a high number of international students are more likely to have better support systems to prevent dropouts.

Sixth, Degree of Exchange with Foreign Universities

It is important to check whether concepts such as “international exchange,” “cooperation,” “partnership,” and “dual degree” show strong connections in the network. If “international exchange” is linked to “support” and “adaptation,” it suggests that universities with active exchange programs may have lower dropout rates among international students. The associated network elements indicate that universities with more foreign university exchanges provide smoother adaptation processes for international students, thereby lowering dropout rates.

Seventh, University’s Capacity for Social Service

It is necessary to verify whether concepts such as “social service,” “community,” “support,” and “volunteer” are connected in the network. If “social service” is linked to “support” and “adaptation,” it suggests that social service activities may help international students settle more effectively. Based on the associated network elements, universities with strong social service programs are likely to support student adaptation, reducing dropout rates.

Based on the analysis results above, the following independent variables are considered meaningful factors influencing the dropout rate of international students:

- University Type (Public vs. Private)
- University Location (Capital Region vs. Non-Capital Region)
- University Admission Quota (Size of Admission Intake)
- Admission Competition Rate
- Number of International Students
- Degree of Exchange with Foreign Universities
- University’s Capacity for Social Service

3. Machine Learning Analysis

1) Variables and Hypothesis

I plan to use the Decision Tree Analysis method among machine learning analysis methods to analyze the characteristics of universities with high dropout rates. As independent variables, I will use the results from the analysis performed above, and I will list and describe the abbreviations of these variables as follows.

Dependent Variable

Dropout rate of international students: The proportion of international students who drop out (dropout)

Independent Variables

Founder type: Whether the university is established by the government or a private entity (1 = public, 2 = private) (founder)

Region: Whether the university is located in the metropolitan area or a non-metropolitan area (1 = metropolitan, 2 = non-metropolitan) (region)

University student enrollment capacity: The total student enrollment capacity of the university (enrollment)

Admission competition rate: The competition rate for new student admissions (competition)

Number of international students: The total number of international students (international)

Level of exchange with foreign universities: The number of students dispatched to foreign universities (exchange)

University’s social service capacity: The ratio of students participating in social service activities to the total number of students (social)

Here’s the translation of the provided text into English:

https://doi.org/10.56294/sctconf20251787

Established Hypotheses

Hypothesis 1: The dropout rate of international students at national universities will be lower than at private universities.

Hypothesis 2: The dropout rate of international students at universities in the capital region will be lower than at universities outside the capital region.

Hypothesis 3: The larger the enrollment capacity of a school, the lower the dropout rate of international students will be.

Hypothesis 4: Universities with higher competition rates for new students will have lower dropout rates among international students.

Hypothesis 5: Universities with a larger number of international students will have lower dropout rates.

Hypothesis 6: Universities with a higher level of international cooperation will have lower dropout rates among international students.

Hypothesis 7: Universities with strong community service capabilities will have lower dropout rates among international students.

2) Results of Decision Tree Analysis

In this study, the Decision Tree Analysis was used to identify factors affecting the dropout rates of international students. The results were interpreted based on the number of observations (N) per node, the percentage of the total data represented by each node (Percent), and the average dropout rate (Mean).

Firstly, the groups with the highest dropout rates were identified as Node 6, Node 12, and Node 10. Among these, Node 6 had the highest average dropout rate at 62,0333 %, which corresponds to a very small percentage (1,4 %, N=3) of the total data, indicating a very high dropout rate at a few universities. Additionally, the average dropout rate at Node 12 was 18,3500 % (N=6), and at Node 10 it was 16,8500 % (N=4), indicating high dropout rates at certain universities under specific conditions.

Next, the groups with moderate dropout rates were Node 3 and Node 9. Node 3 accounted for 10,0 % (N=22) of the total data with an average dropout rate of 15,1500 %. Node 9 accounted for 5,9 % (N=13) of the total data with an average dropout rate of 8,0077 %.

On the other hand, the groups with low dropout rates included Node 4, Node 11, Node 5, and Node 8. Notably, Node 4 accounted for 70,1 % (N=155) of the total data with an average dropout rate of 7,3445 %, suggesting that this group belongs to a cohort of universities with relatively stable retention rates for international students. Moreover, Node 11 had an average dropout rate of 6,4500 % (N=6), and Node 5 had an average dropout rate of 4,7571 % (N=7), both recording relatively low dropout rates. Particularly, Node 8 had an average dropout rate of 0,0000 % (N=5); indicating no dropout cases occurred in this group.

Node	N	Percent	Mean
6	3	1.4%	62.0333
12	6	2.7%	18.3500
10	4	1.8%	16.8500
3	22	10.0%	15.1500
9	13	5.9%	8.0077
4	155	70.1%	7.3445
11	6	2.7%	6.4500
5	7	3.2%	4.7571
8	5	2.3%	.0000

Figure 5. Nodes Derived from Decision Tree Analysis Results

Combining these results suggests that universities with high dropout rates are likely to share certain conditions (e.g., low competition rates, high numbers of international students, non-capital location), necessitating in-depth causal analysis and policy interventions. Meanwhile, identifying the characteristics of universities with low dropout rates and benchmarking their student management strategies may establish more effective policies for retaining international students. Especially, analyzing the operational methods and policy features of universities like Node 8, where no dropout cases occurred, is crucial for deriving practical solutions applicable to other universities.

The figure below represents the results of the Decision Tree Analysis conducted to analyze the influence of

https://doi.org/10.56294/sctconf20251787

independent variables on the dropout rates of international students. It can be seen that all 221 universities are classified into 9 nodes.

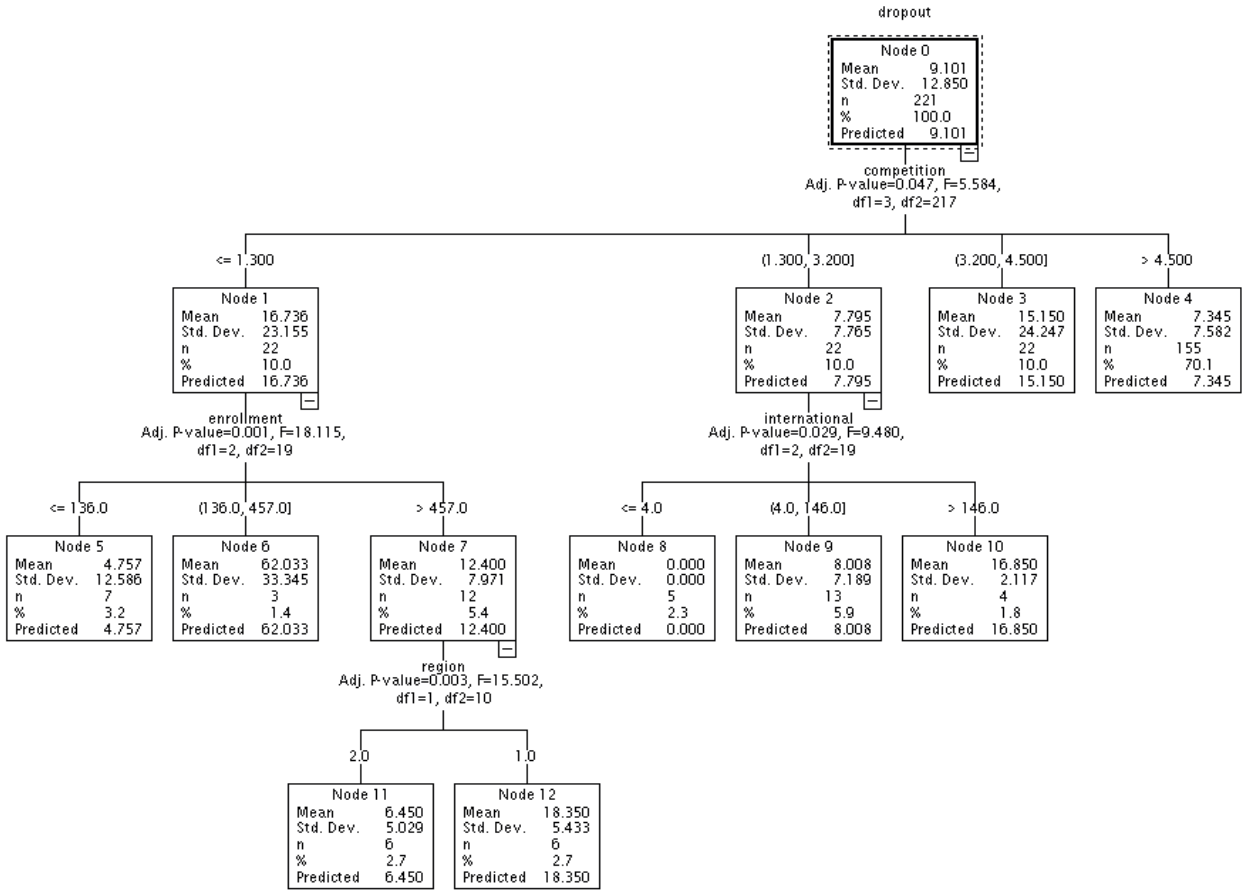

Figure 6. Results of Decision Tree Analysis

The figure above can be described as follows. The analysis results show that the average dropout rate of international students is 10,1 %, with the first major determining factor being the admission competition rate (competition). Based on the competition rate, universities are divided into four groups (less than 3,0; between 3,0 and 3,2; between 3,2 and 4,5; and more than 4,5); and the impact of additional variables on dropout rates was analyzed within each group.

First, for universities with a competition rate (competition) less than 3,0:

The average dropout rate for universities with a competition rate less than 3,0 is 10,79 %, which is slightly higher than the overall average. This group is further branched according to the total student enrollment (enrollment) of the universities, where if the enrollment is less than 457; the dropout rate is low at 4,57%, while it increases to 7,40 % if it exceeds 457.

Second, for universities with a competition rate (competition) between 3,0 and 3,2:

In this case, the average dropout rate is 7,25 %, which is lower than the overall average, and no further division by additional variables occurred, leading to a terminal node.

Third, for universities with a competition rate (competition) between 3,2 and 4,5:

This group recorded a higher average dropout rate of 15,15 % compared to the overall average. Additionally, universities were classified into two subgroups based on the number of international students (international), where the dropout rate was 8,00 % for those with fewer than 140 international students, but sharply increased to 16,85 % for those exceeding 140.

Fourth, for universities with a competition rate (competition) exceeding 4,5:

The average dropout rate in this group is relatively low at 7,35 %. However, it is further branched based on the region (region) where the university is located, with the dropout rate being the lowest at 4,50 % for those in the metropolitan area, while reaching the highest at 18,35 % for those in the non-metropolitan area.

In summary, firstly, the results indicate that the admission competition rate (competition) acts as a major variable. Generally, a higher competition rate tends to result in a lower dropout rate. However, universities

https://doi.org/10.56294/sctconf20251787

with a competition rate between 3,2 and 4,5 exhibit relatively high dropout rates. Secondly, smaller total student enrollments (enrollment) are associated with lower dropout rates. In universities with low competition rates, smaller enrollments tend to reduce dropout rates. Thirdly, a higher number of international students (international) increases dropout rates. In cases where the competition rate is moderate (3,2 to 4,5), having more than 140 international students significantly increases the dropout rate. Fourthly, universities located in metropolitan areas have lower dropout rates. Among universities with a competition rate exceeding 4,5; those located in the metropolitan area have the lowest dropout rate (4,5 %). Conversely, even with the same competition rate, universities in non-metropolitan areas show a significant increase in dropout rates (18,35 %).

Additionally, the figure below represents the results of the Decision Tree Analysis in terms of percentiles, with the x-axis representing the percentile and the y-axis representing the average dropout rate. This visually demonstrates how the predicted dropout rate varies at specific percentiles, clearly showing the differences between universities with high and low dropout rates. Key features of the graph are as follows:

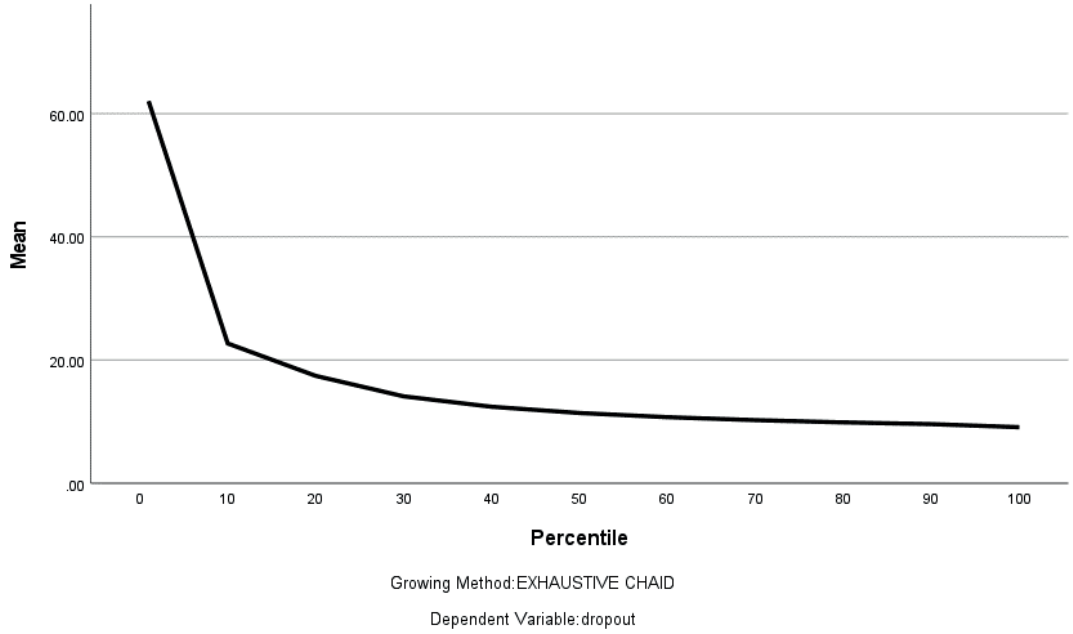

Figure 7. Dropout Rates by Percentile

Firstly, a sharp decrease occurs in the initial range (0-10 %). The average dropout rate for the lowest percentile (i.e., the group with the highest dropout rate) starts at over 60 %. It quickly drops to around 20 % by the 10 % level. This suggests that there are segments with very high dropout rates at some specific universities, and certain characteristics of these universities (such as low competition rates, high numbers of international students, location outside the capital region, etc.) could be major factors causing dropout.

Secondly, a gradual decrease is observed in the 10-50 % range. In the middle percentile range, there is a gradual pattern of decreasing dropout rates. This can be interpreted as the result of independent variables such as admission competition rate (competition), total student capacity (enrollment), and university location (region) having an influence and gradually reducing the dropout rate. That is, the higher the university's admission competition rate, the better managed its student capacity, and the more likely it is located in the capital region, the lower the likelihood of dropout.

Thirdly, a stable pattern is observed after the 50 % range. Once the percentile exceeds 50 %, the average dropout value does not change significantly and maintains a consistent level. This means that in the group with low dropout rates, there are no significant additional fluctuations, and there exists a group of universities that maintain a steady level of dropout. These groups are likely to be universities with a relatively stable international student management system, and are expected to have policy factors that can keep dropout rates consistently low.

CONCLUSIONS

From the analysis of big data, social network analysis, and decision tree analysis on factors affecting the dropout rates of international students, the following findings can be made. Firstly, variables that influence dropout rates include the type of university establishment (national vs. private), the region where the university is located (capital vs. non-capital), the total enrollment size of the university, the size of international students

at the university, the university's community service capability, international exchange capability, and the university's admission competition rate.

Secondly, the hypotheses set in this study were generally accepted. Although the purpose of this study was not hypothesis testing, a tentative presentation of seven hypotheses was made through theoretical discussion, and most of these hypotheses were accepted.

Thirdly, when classifying 221 universities based on dropout rates using decision tree analysis, they were divided into 7 nodes. In particular, the university's admission competition rate (competition) acts as a major variable affecting the dropout rate of international students. Higher competition rates tend to lower the dropout rate. In universities with low competition rates, smaller enrollments tend to decrease the dropout rate. Moreover, the higher the number of international students (international), the higher the dropout rate. Additionally, universities located in the capital region tend to have lower dropout rates.

In conclusion, the admission competition rate, total student capacity, number of international students, and the location of the university are confirmed to be important factors affecting the dropout rate. Based on this, the following policy implications can be drawn. When attracting international students, it is important to maintain a high competition rate, and at the same time, policies managing an appropriate number of international students are needed. Also, as universities in non-capital regions tend to have relatively higher dropout rates, it is necessary to strengthen policies for support and management of international students. Additionally, the smaller the university's capacity, the easier it is to manage international students, which can have a positive impact on reducing dropout rates. As the number of international students increases, the dropout rate tends to increase, so a customized support program for international students is needed considering this.

REFERENCES

1. Kim G. (2024). Mitigating South Korea's Population Decline through Foreign Students.*KJLL, 14*(2), 133-172.

2. Jung J, Kim Y. (2017). Exploring regional and institutional factors of international students' dropout: The South Korea case.*Higher Education Quarterly, 71*(1), 1-18. doi: 10.1111/hequ.12148

3. Choi, M. (2023). *A computational analysis on the role of social relationships in online communication and information diffusion*[Doctoral dissertation, University of Michigan].

4. Choi M, Jurgens D, Romero D. (2023). Analyzing the engagement of social relationships during life event shocks in social media.*Proceedings of the Seventeenth International AAAI Conference on Web and Social Media, 17*, Article 22134. https://doi.org/10.1609/icwsm.v17i1.22134

5. Choi M, Romero DM, Jurgens D. (2023). Profile update: The effects of identity disclosure on network connections and language.*arXiv preprint arXiv:2308.09270*.

6. Beine M, Peri G., Raux M. (2023). International college students' impact on the US skilled labor supply. *Journal of Public Economics, 223*.

7. NAFSA. (2024).*Combating the Enrollment Cliff*.

8. Cambridge Network. (2024).*The Transformative Power of International Education in Local Communities*.

9. upGrad GSP. (2025).*How international students help the local economy*.

10. Chellaraj G, Maskus KE, Mattoo A. (2008). The contribution of international graduate students to U.S. innovation.*Review of International Economics, 16*(3), 444-462.

11. ICEF Monitor. (2025). *UK: Study visa applications and issuances on the rise in 2025*.

12. Shiksha. (2025).*Must-Know Changes for Students in Germany in 2025*.

13. Marginson S. (2013). The impossibility of capitalist markets in higher education.*Higher Education Quarterly, 67*(1), 1-17.

14. Australian Education International. (2019).*Research Snapshot: Jobs supported by international students studying in Australia*.

15. Altbach PG, Knight J. (2007). The internationalization of higher education: Motivations and realities.

Journal of Studies in International Education, 11(3-4), 290-305.

16. OECD. (2022).*Education at a glance 2022: OECD indicators*. Paris: OECD Publishing.

17. Wilkins S, Balakrishnan MS, Huisman J. (2012). Student satisfaction and student perceptions of quality at international branch campuses in the United Arab Emirates.*Journal of Higher Education Policy and Management, 34*(5), 543-556.

18. Andrade MS. (2006). International students in English-speaking universities: Adjustment factors.*Journal of Research in International Education, 5*(2), 131-154.

19. Hawthorne L. (2010). How valuable is "two-step migration"? Labor market outcomes for international student migrants to Australia.*Asia and Pacific Migration Journal, 19*(1), 5-36.

20. Gribble C. (2014).*Employment, work placements & work-integrated learning of international students in Australia*. International Education Association of Australia.

21. Haroon D. (2017).*Python Machine Learning Case Studies: Five Case Studies for the Data Scientist*. Apress.

22. Marr B. (2016).*Big Data in Practice: How 45 Successful Companies Used Big Data Analytics to Deliver Extraordinary Results*. Wiley.

23. Kelleher JD, Mac Namee B., D'Arcy A. (2015).*Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies*. MIT Press.

24. Li K, Jiang H, Yang X. (2023).*Big Data: Algorithms, Analytics, and Applications*. Routledge.

FINANCING

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2022S1A5C2A03092455).

CONFLICT OF INTERESTS

None.

AUTHORSHIP CONTRIBUTION

Conceptualization: Young-Chool Choi, Yun Na Jeong.
Data curation: Young-Chool Choi, Yun Na Jeong.
Formal analysis: Young-Chool Choi, Yun Na Jeong.
Acquisition of funds: Young-Chool Choi, Yun Na Jeong.
Research: Young-Chool Choi, Yun Na Jeong.
Methodology: Young-Chool Choi, Yun Na Jeong.
Project management: Young-Chool Choi, Yun Na Jeong.
Resources: Young-Chool Choi, Yun Na Jeong.
Software: Young-Chool Choi, Yun Na Jeong.
Supervision: Young-Chool Choi, Yun Na Jeong.
Validation: Young-Chool Choi, Yun Na Jeong.
Display: Young-Chool Choi, Yun Na Jeong.
Drafting - original draft: Young-Chool Choi, Yun Na Jeong.
Writing - proofreading and editing: Young-Chool Choi, Yun Na Jeong.