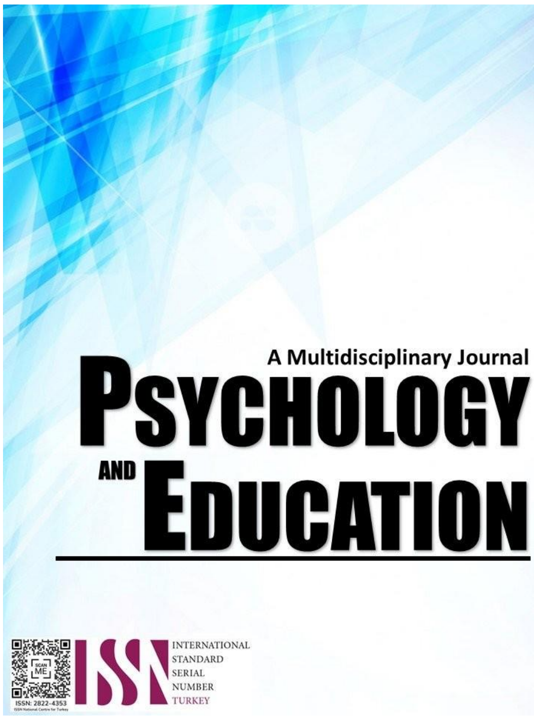
STUDENT MINING USING K-MEANS CLUSTERING: A BASIS FOR IMPROVING HIGHER EDUCATION MARKETING STRATEGIES



PSYCHOLOGY AND EDUCATION: A MULTIDISCIPLINARY JOURNAL

2023
Volume: 14
Pages: 1-8
Document ID: 2023PEMJ1233
DOI: 10.5281/zenodo.8383341
Manuscript Accepted: 2023-27-9

Psych Educ, 2023, **14**: 1-8, Document ID:2023 PEMJ1233, doi:10.5281/zenodo.8383341, ISSN 2822-4353
Research Article

Student Mining Using K-Means Clustering: A Basis for Improving Higher Education Marketing Strategies

Melanie N. Arpay*
For affiliations and correspondence, see the last page.

Abstract

This study aims to enhance marketing strategies in higher education institutions by applying data mining techniques, specifically K-means clustering. The research focuses on Mindanao State University - Lanao del Norte Agricultural College (MSU-LNAC), a tertiary institution in Northern Mindanao, Philippines, with the objective of increasing enrollment. The study utilizes the K-means algorithm to group attributes into different clusters. The clustering analysis provides valuable insights into the characteristics and preferences of the surveyed student population. Based on the findings, recommendations are presented to guide targeted marketing efforts, such as geographic targeting, collaborations with senior high schools, financial assistance programs, and the development of marketing campaigns that emphasize the institution's strengths and advantages. By implementing these recommendations, MSU-LNAC can enhance its recruitment and marketing strategies to attract and retain students effectively.

Keywords: *data mining, higher education institutions, K-Means clustering, marketing strategies*

Introduction

Competition in the higher education sector is forcing the higher educational institutions (HEI) to develop more innovative marketing strategies to face competition and challenges from local schools and international schools all over the world. Mindanao State University - Lanao del Norte Agricultural College (MSU-LNAC), a tertiary school in Northern Mindanao, Philippines, is one institution that recognizes the need to improve its marketing strategy to compete with other schools and to increase its enrollment. Despite its efforts, the college enrollment of MSU-LNAC has remained stagnant over the past 10 years, ranging from 400 to 1000 students. In this research, we aim to explore the potential of data mining techniques, specifically K-Means clustering, to provide a basis for improving MSU-LNAC's marketing strategies thus increasing its enrollment.

To achieve this goal, the institution plans to utilize data mining as it helps them in making proactive and knowledge- driven decisions. Data mining provides many techniques for data analysis. It is a powerful artificial intelligence tool which can discover useful information by analyzing data, categorize information, and summarize relationships. Data mining techniques have gained increasing attention in higher education for their potential to improve academic performance, student services, and marketing strategies.

In this study, we aim to explore the potential of K-Means clustering to gain insights into the characteristics and behaviors of MSU-LNAC's current and prospective student population providing a basis for improving its marketing strategies. By analyzing student data through K-Means clustering, the study will identify groups of students with similar characteristics and behaviors, which can then be used to tailor marketing strategies to the specific needs and preferences of each group.

Institutions could use market segmentation strategies to tailor their marketing messages and recruitment efforts to better attract and retain students who fit into each segment. The findings of this study can have significant implications for higher education institutions seeking to improve their marketing strategies and increase their enrollment. By gaining valuable insights into the profile of students who enroll in their institution, institutions can develop more effective marketing strategies that target the specific needs and preferences of each group. This study contributes to the growing body of literature on data mining in higher education, highlighting the potential of K-Means clustering as a powerful tool for improving marketing strategies and enrollment management practices.

Literature Review

Data mining techniques, particularly K-Means clustering, have gained widespread attention in higher education research. Several studies have explored the application of data mining and K-Means clustering to various aspects of education. [8] provides an overview of data mining applications in education, emphasizing its potential for predicting student performance, identifying at-risk students, and improving educational outcomes. [3] applied clustering analysis to classify student academic performance, demonstrating its effectiveness in grouping students into clusters based on their performance. Utilized K-Means clustering to predict students' academic performance, helping institutions identify students who need additional support.

Proposed an approach that combined K-Means clustering and Decision Tree analysis to improve academic performance by identifying underlying factors contributing to poor performance. Furthermore,conducted a comprehensive review of different methods for determining the optimal number of clusters in K-Means clustering. The authors discussed various techniques, including the elbow method, silhouette method, information criterion approach, and average silhouette width method. These methods are commonly used to evaluate the clustering results and find the most appropriate number of clusters. The elbow method involves plotting the within- cluster sum of squares against the number of clusters and selecting the elbow point where the improvement in clustering starts to diminish. The silhouette method calculates a silhouette coefficient for each data point to assess its cohesion within the cluster. The information criterion approach employs statistical measures like the Bayesian Information Criterion (BIC) or Akaike's Information Criterion (AIC) to determine the optimal number of clusters based on model fit. The average silhouette width method evaluates the compactness and separation of clusters using the average silhouette coefficient across all data points. By comparing the strengths and limitations of these methods, researchers can choose the most suitable approach for determining the number of clusters in their specific educational context.

The methodology employed in these studies typically involved data collection, preprocessing, and analysis. Collected student academic performance data from a Malaysian university and performed data preparation to make it suitable for analysis. They applied the K-Means clustering algorithm to group students based on their academic performance, followed by the evaluation of clustering results using external validation indices such as the silhouette index and Dunn index. Similarly, aimed to classify and predict students' GPA by utilizing the K-Means clustering algorithm. They employed a three-step methodology consisting of data preprocessing, K-Means clustering, and prediction modeling. The authors utilized a dataset of student admission records and academic performance data for analysis. The clustering results helped group students based on their admission data and academic performance, contributing to the student admission process. The predictive modeling techniques accurately predicted students' GPA based on their admission data and cluster assignment.

Used an adaptive learning system to collect data on middle school students and applied statistical methods, including principal component analysis (PCA) and K-Means clustering, to analyze the collected data and generate student profiles. The adaptive learning system tracked student interactions with online learning materials, and PCA was used to reduce the dimensionality of the data. K-Means clustering was then applied to identify groups of students with similar learning behaviors and preferences. The study demonstrated the potential of adaptive learning systems and student profiling approaches to enhance personalized learning experiences and improve student academic performance.

Furthermore, conducted a survey to explore the application of market segmentation theory in understanding student behavior when selecting a school or department. They used cluster analysis to identify four distinct segments of students based on their preferences and decision-making factors. The study revealed significant differences in demographic characteristics and attitudes among the identified segments. This information can assist educational institutions in tailoring their marketing strategies to attract students from different segments effectively. These studies highlighted the efficacy of data mining techniques, particularly K-Means clustering, in various educational contexts. The clustering results provided valuable insights into student performance, admission processes, decision-making behavior, and factors affecting academic outcomes. By employing appropriate methodologies encompassing data collection, preprocessing, and analysis, researchers were able to uncover meaningful patterns and develop models for prediction and decision support. The utilization of different methods for determining the optimal number of clusters, as discussed by [11], further enhanced the accuracy and effectiveness of the clustering results. These findings collectively underscored the potential of data mining and K-Means clustering in improving educational practices, fostering student success, and facilitating evidence-based decision-making in higher education.

Methodology

Data Collection

The survey questionnaire was drafted based on related literature, consultation from experts, and deliberation from stakeholder. The questionnaire includes various attributes such as age, sex, religion, address, senior high school graduated category/classification, senior high school graduated, monthly family income, education level of parents, occupation of parents, senior high school academic achievement, senior high school extracurricular activities or involvement, degree program taken, year level, decision maker in enrolling in MSU-LNAC, preferred degree program, reasons for enrolling in MSU-LNAC, and sources of information about MSU-LNAC. The dataset from this study was collected from the entire college students of MSU-LNAC. The researcher conducted a school-wide data survey to get a comprehensive view of students' data.

Integration and Conversion

The collected data was integrated into a single dataset using Microsoft Excel. Most machine learning algorithms generally require numeric input and output variable [3]. To achieve this, the technique of ordinal encoding was employed. It involved assigning numerical representations to the different categories within each variable. This transformation enabled the subsequent machine learning algorithm, specifically K-means clustering, to process the data effectively. The conversion process was performed using Microsoft Excel.

Data Preprocessing

Python programming in Google Colaboratory was utilized for data preprocessing and clustering. In the preprocessing phase, the dataset underwent various steps to ensure its quality and suitability for analysis. This included cleaning and feature selection. Data cleaning is an essential step in handling missing values within a dataset. In this study, the researcher employed the fillna() method to replace missing values. This approach is appropriate for the survey questionnaire utilized in the study because it involved ranking questions where students had the option to skip choices that were not applicable to them. By filling the missing values with zeros, it allows for a consistent representation of the data, ensuring that all attributes have values assigned to them. This strategy helps to maintain the integrity and completeness of the dataset, enabling subsequent analysis and processing to be performed accurately.

In this study, Domain knowledge and Variance threshold selection technique were used in feature selection. Domain knowledge includes leveraging the expertise and understanding of the domain in which the dataset belongs to identify and select relevant features. By incorporating domain knowledge, the feature selection process can be more targeted and informed. In this technique, attributes that were identified irrelevant for clustering were removed in this phase. Variance threshold selection technique on the other hand, is an unsupervised feature selection technique [3] that discards attributes that have low variability and carry little discriminative information. Features that were identified that have low-variance were analyzed first before completely removing in the dataset.

The preprocessing phase is crucial before clustering to improve the efficiency, interpretability, and accuracy of the clustering analysis. By selecting the most informative features, we can uncover meaningful patterns and gain valuable insights from the clustering results.

Clustering

Clustering is a fundamental technique in unsupervised machine learning used to discover patterns and group similar data points together [4] [10]. In this research, the goal of clustering is to identify distinct groups or clusters of students enrolled in MSU-LNAC based on their preferences and characteristics, which can then be used to inform marketing strategies to increase enrollment.

Before applying the K-means clustering algorithm, the elbow method was employed to determine the optimal number of clusters (k) for the dataset [11]. It works by calculating the inertia, which is the sum of squared distances between each data point and its nearest cluster centroid. The method involves running the K-means algorithm for different values of k, the number of clusters, and computing the inertia for each case. The results are then plotted on a graph, with k on the x-axis and inertia on the y-axis. The elbow point on the graph represents a bend or elbow in the curve, where the inertia starts to flatten out. This point indicates the optimal number of clusters [11]. The elbow method visually assists in determining the appropriate number of clusters before applying the K-means algorithm. Finally, K-means clustering, the most widely used clustering algorithm in data mining, was utilized with the determined number of clusters to

Psych Educ, 2023, **14**: 1-8, Document ID:2023 PEMJ1233, doi:10.5281/zenodo.8383341, ISSN 2822-4353
Research Article

cluster the data and reveal meaningful patterns among the enrolled students in MSU-LNAC.

K-means clustering is a popular unsupervised learning algorithm utilized in this study. It is commonly applied when dealing with unlabeled data [12]. The algorithm employs a distance-based approach to assign objects to clusters based on their similarity. The number of clusters is determined by specifying the value of k. In this study, the optimal number of clusters determined through the elbow method. The K-means algorithm is shown below.

Algorithm 1 Basic K-means Algorithm	
1:	Specify the number of k clusters to assign
2:	Randomly initialize k centroids
3:	Repeat
4:	Form k clusters by assigning all points to the closest centroid
5:	Recompute the centroid of each cluster
6:	Until the centroids don't change

Table1. *Traditional K-means algorithm*

Analysis

The analysis phase is a crucial step in interpreting and deriving insights from the results obtained through clustering. The main objective is to gain a deeper understanding of the distinct groups or clusters of students enrolled in MSU- LNAC based on their preferences and characteristics. After performing the clustering, the clustered data was thoroughly analyzed and interpreted. Initially, the mean values of each attribute in each cluster were calculated. However, since the dataset primarily consisted of categorical variables, it became challenging to analyze the similarity within each cluster and distinguish it from other clusters based solely on the means. As a result, a more thoughtful approach was adopted, leveraging domain knowledge and careful planning.

To facilitate the analysis, the collected data was categorized into specific categories, such as demographics, senior high school background, socio-economic status, decision-making insights, and sources of information. By grouping the attributes within these categories, it becomes possible to explore the characteristics and preferences that may be relevant for each cluster. A variable containing categories that lack a natural order or ranking is referred to as a nominal scale. Calculations like mean, median, or standard deviation would be pointless for nominal variable because they are arbitrary. Thus, in analyzing the data, a shift from mean values to frequency counts was implemented. Frequency count is a descriptive statistics and data analysis technique that provides insights into the distribution of data and identifies the most common or rare categories within each cluster. This approach allows researchers to determine the prevalence of specific attributes or preferences within each group, providing a more comprehensive understanding of their characteristics.

By utilizing frequency counts, researchers can identify dominant trends, patterns, or preferences within each cluster. This information is valuable in formulating targeted strategies and interventions tailored to the unique needs and preferences of each group. The utilization of frequency counts, combined with the categorization of attributes, strengthens the analysis and provides a more nuanced understanding of the characteristics and preferences of the student population at MSU-LNAC.

Results and Discussion

Out of the 734 college population of MSU-LNAC for S.Y. 2022-2023, 404 (55%) students were able to answer since graduating students were deployed for internship and on-the- job training, and some students were absent during the conduct of the survey.

The Preprocessing stage handles all the missing values by replacing 0s through fillna(0) method. This also included the categorizing of attributes based on the target outcome as shown below.

Table 2. *List of Target output and Cluster Attributes*

Category	*Attribute*
Demographic Information	Age Sex Religion Permanent Resident Address
Senior High School Background	Senior High School graduated category/classification Senior High School graduated from Entry qualification SHS Academic achievement
Socio-economic Factors	Monthly family income Education level of parents Occupation of parents
Decision-Making Factors	Decision maker in enrolling in MSU-LNAC Reasons for enrolling in MSU-LNAC
Sources of Information	Sources of information MSU-LNAC School to School Campaign Family/Relatives Peers/ Friends Radio Advertisement Social Media Advertisement Others:

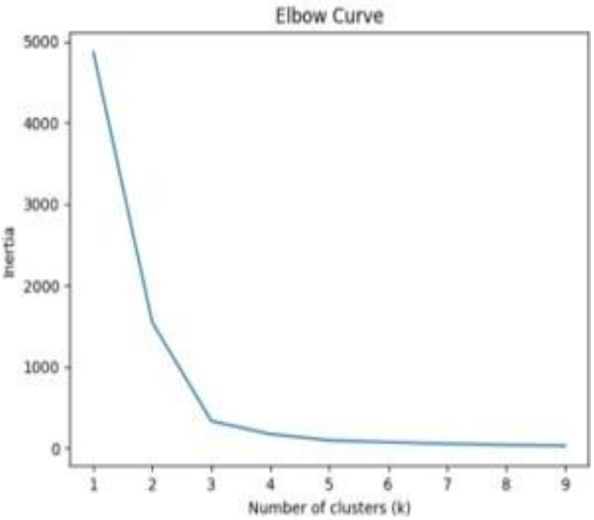


Fig 1. *Determine the number of Clusters (K) through Elbow method*

Figure 1 illustrates the application of the Elbow method to determine the optimal number of clusters for the demographic information category. The graph displays the number of clusters on the x-axis and the corresponding distortion or inertia on the y-axis. The graph exhibits a distinct elbow at three clusters, indicating that the optimal number of clusters for the demographic information category is three.

The same methodology was applied to determine the optimal number of clusters for the senior high school background, socio-economic factors, decision-making factors, and sources of information categories.

The table 3 shows the clustering result based on demographic information attributes using the K-means algorithm with 3 clusters. Cluster 1 (17%) represents a group of students originating from Kapatagan, Lala, Tubod, and Baroy in Lanao del Norte. These locations are approximately 36-43 kilometers away from MSU-LNAC. It is observed that the majority of students in this cluster identify themselves as Christians. Cluster 2 (6%) consisted of students from areas outside of Lanao del Norte and Lanao del Sur. These students were predominantly Christians. Cluster 3 (77%) encompassed students from Salvador and Sapad municipalities, approximately 44 kilometers away from MSU-LNAC, as well as Sultan Naga Dimaporo, where MSU-LNAC is situated. In this cluster, both Christian and Muslim students were represented. I observed that a significant proportion of students in this cluster were graduates from within SND Municipality and predominantly identified as Muslims.

Table 3. *Demographics Cluster Result*

	Attributes	*Address*	*Christianity*	*Islam*	*Others*	*Cluster 1*	*Cluster 2*	*Cluster 0*
1	a) Kapatagan, Lanao del Norte	22	21	0	1	22	0	0
2	b) Lala, Lanao del Norte	39	36	0	3	39	0	0
3	c) Tubod, Lanao del Norte	7	7	0	0	7	0	0
4	d) Baroy, Lanao del Norte	1	0	1	0	1	0	0
5	e) Maigo, Lanao del Norte	0	0	0	0	0	0	0
6	f) Salvador, Lanao del Norte	15	7	7	1	0	0	15
7	g) Sapad, Lanao del Norte	9	7	2	0	0	0	9
8	h) Nunungan, Lanao del Norte	2	0	2	0	0	0	2
9	i) SND, Lanao del Norte	260	116	137	7	0	0	260
10	j) Picong, Lanao del Sur	22	0	22	0	0	0	22
11	k) Malabang, Lanao del Sur	1	1	0	0	0	0	1
12	l) Balabagan, Lanao del Sur	0	0	0	0	0	0	0
13	m) Calanogas, Lanad del Sur	0	0	0	0	0	0	0
14	n) Tukuran, Zamboanga del Sur	0	0	0	0	0	0	0
15	o) Aurora, Zamboanga del Sur	1	1	0	0	0	1	0
16	p) Others:	25	17	8	0	0	25	0

These findings highlight the relationship between students' addresses and their religious affiliation, with Cluster 1 and 2 comprising predominantly Christian students from areas outside Sultan Naga Dimaporo, and Cluster 3 demonstrating a mix of Christian and Muslim students, with a larger proportion originating from within SND Municipality.

Table 4. *Senior High School Background Cluster Result*

	Senior High School Graduated Attributes	*SHS_Graduated*	*Entry Qualification SASE*	*Entry Qualification CET*	*Entry Qualification CBP*	*Cluster 1*	*Cluster 2*	*Cluster 3*
1	a) Sultan Naga Dimaporo Memorial Integrated School (SND MIS)	53	39	7	7	53	0	0
2	b) MSU-LNAC Senior High School	85	71	0	14	85	0	0
3	c) Andres Bersales National High School (ABNHS)	12	9	3	0	12	0	0
4	d) Sultan Ali Dimaporo Integrated School (SADMIS)	43	32	4	7	43	0	0
5	e) Placida Mequibas National High School (PMNHS)	1	1	0	0	0	1	0
6	f) Bansarvil National High School (BNHS)	35	21	7	7	0	35	0
7	g) Kapatagan, National High School (KNHS)	13	10	3	0	0	13	0
8	h) Lala National High School (LNHS)	45	27	10	8	0	45	0
9	i) Lanao del Norte National Comprehensive High School (LNNCHS)	4	3	0	1	0	4	0
10	j) Salvador National High School	13	11	2	0	0	13	0
11	k) Panoloon National High School	4	1	2	1	0	4	0
12	l) Christ the King College de Maranding (CKCM)	2	1	1	0	0	0	2
13	m) North Central Mindanao Colleges (NCMC)	13	8	1	4	0	0	13
14	n) Lanao School of Science and Technology Incorporated (LSSTI)	1	0	1	0	0	0	1
15	o) ICI Kapatagan	4	4	0	0	0	0	4
16	p) Others:	76	45	22	9	0	0	76
		404	283	63	58	193	115	96

Psych Educ, 2023, **14**: 1-8, Document ID:2023 PEMJ1233, doi:10.5281/zenodo.8383341, ISSN 2822-4353
Research Article

The Senior High School background classification involved clustering the data into 3 clusters based on SHS graduation category, SHS graduated from, and entry qualifications.

Cluster 1: Primarily consists of students who graduated from Public Senior High Schools within Sultan Naga Dimaporo (SND), with 41% of them passing the SASE and CET qualifications.
Cluster 2: Comprises students from public schools including Bansarvil National High School, Kapatagan NHS, Lala NHS, and Salvador NHS. Among them, 24% passed the SASE and CET entry qualifications.
Cluster 3: Includes students from private schools outside of SND, as well as those from unmentioned schools. In this cluster, 20% of students passed the SASE and CET qualifications.

Additionally, 15% of students belong to the College Bound Program, a bridging program for those who did not meet the SASE/CET cut-off scores.

These findings shed light on the SHS backgrounds of the students, with distinct clusters representing students from public schools within SND, public schools outside SND, and private schools. Additionally, the analysis highlights the varying proportions of students who have met the SASE and CET entry qualifications across the clusters.

Table 5. *Socio-Economic Status Cluster Result*

	Occupation of Parents	*Cluster 0*	*Cluster 1*
1	a) Farmers	0	230
2	b) Fishermen	0	32
3	c) Business Owner	0	21
4	d) Teachers and Education Professionals	0	14
5	e) Healthcare workers (Doctor, Nurse, BHW)	0	3
6	f) LGU and National government agencies workers	12	0
7	g) Engineers	0	0
8	h) Technicians	3	0
9	i) Service Workers (Hotel and restaurant staff, Housekeepers, Caregivers)	29	0
10	j) Drivers	33	0
11	k) Others:	0	0
	Monthly Family Income	*Cluster 0*	*Cluster 1*
1	a) Less than ₱9,100	88	251
2	b) Between ₱9,100 to ₱18,200	10	31
3	c) Between ₱18,200 to ₱36,400	5	13
4	d) Between ₱36,400 to ₱63,700	1	1
5	e) Between ₱63,700 to ₱109,200	0	2
6	f) Between ₱109,200 to ₱182,000	0	0
7	g) At least ₱182,000 and up	0	2

Based on table 5, Cluster 0 has a more diverse range of occupations among parents, while Cluster 1 is predominantly composed of parents working in agricultural occupations, such as farming and fishing. Additionally, Cluster 0 includes students from a wider range of income brackets, including higher income levels, compared to Cluster 1, which primarily consists of students from lower-income brackets.

Table 6. *Decision Making Factors Cluster Result*

Decision making Factors	*Cluster 1*	*Cluster 2*
Parent's Decision	16	
Scholarships	15	
Affordability/Costs	14	
School's Image/ Reputation/Name	13	
Distance of school from home/location	12	
Family members also are schooling/graduated here	11	
Quality Education	10	
Transportation/Convenience of school location		9
Career/Employment Opportunities		8
Peace and Security		7
School's Extra-Curricular Activities/ Social Life		6
Religious Beliefs		5
Facilities and Equipment		4
Information Sources/Advertising		3
Friends/Peers		2
Size of Library Collection		2

The analysis of the clusters reveals that individuals in Cluster 1 consider factors such as parent's decision, financial aspects, school reputation, location convenience, and family influence when choosing a school. They also value quality education, career opportunities, and extracurricular activities. In contrast, individuals in Cluster 2 prioritize factors like transportation convenience, career prospects, peace and security, extracurricular activities, and social life of the school. Understanding these clusters helps identify the different factors that influence school choice, with Cluster 1 emphasizing parental and financial aspects, while Cluster 2 prioritizes practical considerations and overall school experience.

Table 7. *Sources Of Information Cluster Result*

Sources of Information	*Cluster 0*	*Cluster 1*	*Cluster2*
MSU-LNAC School to School Campaign	17	29	30
Family/ Relatives	32	132	57
Peers/ Friends	9	48	17
Radio Advertisement	1	0	2
Social Media Advertisement	1	4	7

Psych Educ, 2023, **14**: 1-8, Document ID:2023 PEMJ1233, doi:10.5281/zenodo.8383341, ISSN 2822-4353
Research Article

Table 7 concludes that Cluster 1 relies heavily on Family/Relatives and Peers/Friends as sources of information. Cluster 2 has a higher emphasis on the MSU- LNAC School to School Campaign, while Cluster 0 shows a relatively lower presence of all the sources of information.

Discussion

The clustering analysis of the surveyed students at MSU- LNAC has provided valuable insights into their characteristics and decision-making factors. The analysis reveals distinct profiles within different clusters based on demographics, senior high school backgrounds, socio- economic status, decision-making factors, and sources of information. These findings emphasize the correlation between students' addresses and religious affiliation, the varying academic achievements across clusters, the influence of geographic location on senior high school backgrounds, the impact of socio-economic factors on enrollment decisions, and the distinct preferences guiding students' school choices. These findings collectively contribute to a deeper understanding of the factors influencing student enrollment and can inform targeted recruitment and marketing strategies for MSU-LNAC.

Based on the findings of the study, several recommendations have been formulated to enhance the marketing strategies of Mindanao State University - Lanao del Norte Agricultural College (MSU-LNAC) and effectively attract and retain students. The recommendations are as follows:

Targeted Geographic Marketing: Develop marketing materials that showcase the university's commitment to promoting interfaith harmony and cultural diversity, emphasizing the availability of facilities, programs, and support services that cater to students from different religious backgrounds. For areas outside SND, target outreach efforts by organizing information sessions, school visits, and online campaigns. Collaborate with Senior High Schools: Strengthen collaborations with Senior High schools within SND by offering targeted information sessions and specialized support programs. Showcase success stories of students who transitioned from SND schools to MSU-LNAC, highlighting the smooth academic integration and career opportunities available to them. Develop partnerships with private senior high schools and expand outreach efforts to reach students from unmentioned schools. Highlight the advantages of choosing MSU-LNAC as the next educational step after private senior high schools, emphasizing the university's academic reputation, strong faculty, safe environment, and extensive support services.

Socio-economic Status: Since most of the students fall below 9,100 monthly income, develop targeted financial assistance programs and scholarships specifically aimed at supporting students from lower-income backgrounds. Highlight the affordability of education at MSU-LNAC, the availability of work-study programs, and the university's commitment to providing equal educational opportunities. Decision-making Factors: Create marketing campaigns that highlight MSU-LNAC's strong reputation, affordability, and the positive impact parents can have on their child's education. Provide testimonials from current parents highlighting their satisfaction with the university and the positive outcomes for their children. Find strategies to ensure parents' presence during the school-to-school campaign since parents' decisions significantly influence enrollment in MSU- LNAC. Additionally, develop targeted campaigns that showcase MSU-LNAC's convenient transportation options, career development programs, safe campus environment, vibrant extracurricular activities, and strong sense of community. Highlight success stories of students who have benefited from these factors to demonstrate their value.

References

A.F.M Nafuri, N.S Sani, N.F.A Zainudin, A.H.A Rahman, M. Aliff, “Clustering Analysis for Classifying Student Academic Performance in Higher Education,” https://www.mdpi.com/journal/applsci, 2022

A. Algarni, “Data Mining in Education , ” International Journal of Advanced Computer Science and Applications, Vol. 7, No. 6, pp 456 – 461, 2016

A. Zhang, O. Walker, K. Nguyen, Dai, J. Dai, A. Chen, M. Lee, “Deliberating with AI: Improving Decision-Making for the Future through Participatory AI Design and Stakeholder Deliberation,” Proc. ACM Hum.-Comput. Interact., Vol. 7, No. CSCW1, Article 125, pp 125-125:33, 2023

I. Maniu and G.C. Maniu, “Educational Marketing: Factors influencing the selection of a University, ” SEA – Practical Application Science,Vol. 2, No. 3 (5), pp 37-41, 2014

J.T Lalis, “A New Multiclass Classification Method for Objects with Geometric Attributes Using SimpeLinear Regression, ” International Journal of Computer Science , 43:2, IJCS_43_2_08, 2016

O. Oyelade, O. Oladipupo, I. Obagbuwa, “Application of K-Means Clustering algorithm for prediction of Students’ Academic Performance , ” International Journal of Computer Science and Information Security, Vol. 7, No. 1, pp 292 - 295, 2010

R. Santosa, Y. Lukito, A.R. Chrismanto, “Classification and Prediction of Students’ GPA Using KMeans Clustering Algorithm to Assist Student Admission Process,” Journal of Information Systems Engineering and Business Intelligence, Vol. 7, No. 1, pp 1-10, 2021

S. Wang, M. Fenf, M. Bienkowski, C. Christensen, W. Cui, “Learning from an Adaptive Learning System: Student Profiling among Middle School Students,” Proceedings of the 11th International Conference on Computer Supported Education (2019)

T. Kodinariya, P. Makwana, “Review on determining number of Cluster in K-Means Clustering, ” International Journal of Advance Research in Computer Science and Management Studies, IJARCSMS, Vol. 1, No. 6, pp 90 – 95, 2013.

Y.F Chen and C.H. Hsiao, “Applying market segmentation theory to student behavior in selecting a school or department , ” New Horizons in Education, Vol. 57, No. 2, pp 32-43, 2009

Affiliations and Corresponding Information

Melanie N. Arpay
Mindanao State University
Lanao Del Norte Agricultural College – Philippines