Available online at website: https://jurnal.iaii.or.id/index.php/RESTI

JURNAL RESTI

(Rekayasa Sistem dan Teknologi Informasi)

Vol. 9 No. 3 (2025) 659 - 669 e-ISSN: 2580-0760

A New Framework for Dynamic Educational Marketing Segmentation in Student Recruitment: Optimizing Fuzzy C-Means with Metaheuristic Techniques

Rizal Bakri[1]*, Bobur Sobirov[2], Niken Probondani Astuti[3], Ansari Saleh Ahmar[4], Pawan Kumar Singh[5]

[1]Department of Digital Business, Faculty of Economics and Business, Makassar State University, Makassar, Indonesia

[2]Department of Economics, Samarkand Branch of Tashkent State University of Economics, Uzbekistan

[1,3]Statistics Research Group, STIEM Bongaya, Makassar, Indonesia

[4]Department of Economic, Financial and Actuarial Mathematics, Universität de Barcelona, Spain

[5]Department of Economic, Lakshmibai College, University of Delhi, India

[1]rizal.bakri@unm.ac.id, [2]mrbobursobirov@gmail.com, [3]niken.probondani@stiem-bongaya.ac.id, [4]aahmarah9@alumnes.ub.edu, [5]pawankumarsingh@lb.du.ac.in

Abstract

An effective educational marketing strategy requires accurate school segmentation to enhance new student recruitment. Traditional segmentation methods such as K-means are often used, but they have limitations in capturing the flexibility of school characteristics. Fuzzy C-Means (FCM) offers a more adaptive approach by allowing each school to simultaneously have a degree of membership in several clusters. However, the performance of FCM highly depends on determining parameters such as the number of clusters (k) and the level of fuzziness (m), which are not always optimal when determined manually. This study develops a new framework for dynamic educational marketing segmentation in student recruitment by optimizing FCM using three metaheuristic techniques: Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and Differential Evolution (DE). Performance was evaluated using the Fuzzy Silhouette Index (FSI). The experimental results showed that DE yielded the best results with the highest FSI value (0.8023), producing eight main clusters based on the Recency, Frequency, and Monetary (RFM) model. Based on the clustering results, a personalized and adaptive marketing strategy was designed to enhance the effectiveness of student recruitment. The proposed framework enhances segmentation accuracy and supports the implementation of dynamic data-driven marketing in the context of higher education. This study also opens new directions for educational data mining research and machine-learning-based marketing strategies.

Keywords: dynamic educational marketing; fuzzy C-Means; metaheuristic optimization; RFM; student recruitment

How to Cite: R. Bakri, B. Sobirov, N. P. Astuti, A. S. Ahmar, and P. K. Singh, "A New Framework for Dynamic Educational Marketing Segmentation in Student Recruitment: Optimizing Fuzzy C-Means with Metaheuristic Techniques", *J. RESTI (Rekayasa Sist. Teknol. Inf.)* , vol. 9, no. 3, pp. 659 - 669, Jun. 2025.

*Permalink/*DOI: https://doi.org/10.29207/resti.v9i3.6515

Received: April 6, 2025

Accepted: June 11, 2025

Available Online: June 22, 2025

1. Introduction

Educational marketing has become a key factor in determining the success of student recruitment at universities, especially private universities that heavily rely on the number of new students each year. In this context, an effective marketing approach not only focuses on institutional branding but also on segmentation strategies that can identify potential schools as the main source of prospective students [1]. Without effective segmentation, institutions risk missing out on reaching diverse student populations, which can significantly impact enrollment numbers. Previous studies in educational marketing and data-driven recruitment have applied clustering methods to identify and group target schools or student profiles based on various features. For instance, research by [2] conducted a comprehensive school segmentation based on key attributes such as academic quality, student population, financial background, and geographical proximity, providing valuable insights for higher education institutions striving to refine their recruitment strategies. Similarly, [3] employed geographic attributes specifically to advance spatial school segmentation. Collectively, these studies demonstrated that effective segmentation models can personalize marketing messages and enhance enrollment efficacy. Nonetheless, many of these studies depend on static clustering techniques that lack the necessary flexibility in marketing applications.

Various clustering methods have been used in education market segmentation for student recruitment, one of which is K-Means clustering, which is the most popular clustering technique used in various studies [4], [5] and based on the Recency, Frequency, and Monetary (RFM) model [6], [7]. However, the segmentation using K-Means method in existing research has a major limitation, which is that each object can only fit into one cluster. This does not reflect the reality in the field when conducting promotions, where a school can have overlapping characteristics with several different clusters (hard clustering) [8]. In addition, K-means is highly dependent on the selection of the number of clusters (k) that must be predefined, so it can produce suboptimal segmentation if the value of k is not chosen properly [9]. This limitation underscores the need for more sophisticated approaches that capture the complexity of educational marketing dynamics.

To address these limitations, Fuzzy C-Means (FCM) has emerged as a more flexible clustering method, probability-based segmentation [10]. Unlike K-Means, FCM allows a school to have membership in more than one cluster [11], making it more realistic in representing the variation of school characteristics in education marketing [12]. FCM helps to perform dynamic segmentation [13] and has been adopted in various Educational Data Mining (EDM) studies [12], [14]-[17]. However, the application of FCM still has a major challenge, which is the selection of optimal parameters, specifically the number of clusters (k) and the fuzziness level (m) [18]. Without the right parameters, FCM segmentation results can be unstable and less accurate [19], which in turn can reduce the effectiveness of education marketing strategies. This instability emphasizes the importance of refining the parameter selection process to enhance overall segmentation quality.

To solve this problem, several studies have used metaheuristic techniques to optimize the parameters of FCM. For instance, a study by [20] showed that FCM with the Particle Swarm Optimization (PSO) algorithm was able to improve stability and accuracy. Another study by [21] demonstrated that combining FCM with a Genetic Algorithm (GA) provided a more versatile and adaptive solution in large-scale data analysis. Similarly, [22] used Differential Evolution (DE) to optimize FCM and reported high classification accuracy. Additional comparisons by [23] and [24] showed that hybrid FCM approaches with metaheuristic algorithms such as PSO, GA, Firefly Algorithm (FA), and Artificial Bee Colony (ABC) outperformed traditional clustering methods. Furthermore, [25] confirmed that DE showed better performance than PSO and GA in FCM optimization. Although these methods have proven effective in various domains, their application in educational data mining remains limited, highlighting a gap in the literature that this research aims to address.

This research aims to develop a new framework for education marketing segmentation by optimizing FCM using GA, PSO, and DE to fill this gap. Specifically, this research seeks to determine the optimal combination of k and m parameters, compare the performance of the three optimization methods to improve segmentation accuracy, and evaluate their impact on student recruitment strategies in higher education. The novelty of this research is the integration of the FCM method with metaheuristic-based optimization in the context of education marketing, which is still very limited in previous studies. In addition, this research also provides various marketing strategy techniques based on the results of school clustering, ensuring that educational institutions can implement data-driven strategies effectively.

In summary, the contribution of this study lies in offering a more flexible and realistic segmentation framework for educational institutions by integrating FCM with metaheuristic optimization, improving clustering accuracy, and providing actionable insights for targeted and data-driven student recruitment strategies. This innovative approach not only enhances the effectiveness of marketing strategies but also empowers institutions to engage with prospective students more successfully.

2. Methods

The method used consists of several main stages, namely data collection and preparation, RFM model calculation, segmentation using FCM, FCM parameter optimization, evaluation of results with Fuzzy Silhouette Index (FSI), and designing data-based dynamic marketing strategies.

2.1 Data collection and preparation

This study utilizes data obtained from the database of Student Admission Information System of STIEM Bongaya for the period 2020–2024. The dataset contains 2,342 student records originating from 460 different schools. The raw data includes three primary variables: registration date, student name, and school ID. The data collection process involved retrieving records from the database system as a .csv file. The export was ensured to include all relevant records while maintaining the integrity of the data.

In the data preparation stage, the student-level data were aggregated by school to create a more structured and school-centric format. The resulting Table 1 includes the date of the registration, the school's ID, and the total number of students who enrolled. Several preprocessing steps were conducted to ensure data quality, including removing duplicate entries, handling missing values, and standardizing date formats.

The prepared dataset was then used to construct the Recency, Frequency, and Monetary (RFM) model, which serves as the foundation for the subsequent clustering analysis. All data processing and transformation steps were performed using R with the dplyr [26] and reshape2 [27] packages.

Table 1. Aggregating of student registration dataset

Registration Date	School's ID	Number of students
01-02-2020	S001	4
12-04-2021	S078	2
20-06-2022	S125	3
23-03-2023	S367	1
…	…	…
26-06-2024	S460	3

2.2 Create RFM model

In this step, recency, frequency, and monetary (RFM) calculations are performed for each school to measure their level of contribution in new student enrollment. RFM is a commonly used technique in customer segmentation, and in the context of this research, it is used to group schools based on student enrollment patterns at STIEM Bongaya University. Recency (R) measures how recently the school last sent students to enroll. The more recently the school contributed, the higher the score. This metric is crucial as it indicates the immediacy of a school's engagement with the university, thereby reflecting its current relevance in recruitment efforts. Frequency (F) measures how often the school sends students in the analyzed period (2020-2024). Schools that consistently send students more frequently get a higher score. This aspect assesses the loyalty and ongoing relationship of the school with the university, serving as an indicator of a school's established engagement in student recruitment. Monetary (M) refers to the total number of students sent by the school in the period. Schools with a larger number of applicants get a higher score. This measure highlights the overall contribution of the school to the university's enrollment, with a higher number suggesting greater effectiveness in attracting students.

Each RFM variable is assigned a value scale of 1-5, where a score of 1 indicates the lowest contribution and a score of 5 the highest contribution. The scale is determined based on the quantile distribution of the data, enabling an equitable assessment across varying scales of school performance. Specifically, the scores of schools are ranked for each RFM metric, and scores are assigned based on their percentile rankings within the dataset. For example, the top 20% of schools in terms of recency would receive a score of 5, the next 20% a score of 4, and so on, with the lowest-performing schools receiving a score of 1. Because the RFM scale is consistent across the three dimensions, no further normalization is required in the clustering process. This consistency promotes the efficacy of clustering methods by ensuring that each dimension contributes equally to the analysis. The package used to create the RFM dataset is the rfm package [28], which facilitates the efficient computation of RFM metrics and allows for the seamless integration of these metrics into the subsequent clustering analysis.

2.3. Fuzzy C-Means (FCM) clustering

After the recency (R), frequency (F), and monetary (M) values are calculated for each school, the next step is to perform segmentation using the fuzzy C-means (FCM) method. FCM is a probability-based clustering algorithm [29] that provides a flexible cluster assignment, allowing one school to have membership degrees in more than one cluster. This enables a nuanced approach in which schools are not rigidly classified into exclusive groups but are assigned probabilities of membership across multiple clusters, reflecting the complexity of real-world student recruitment scenarios.

This algorithm works by minimizing an objective function such as Formula 1.

$$J_m = \sum_{i=1}^{N} \sum_{j=1}^{k} u_{ij}^{m} \left|\left|x_i - c_j\right|\right|^2 \quad (1)$$

Where J_m is objective function that must be minimized, N is a number of data, k is a number of cluster, x_i is a the i-th data vector, c_j is the center of the j-th cluster, u_{ij} is fuzzy membership degree, and m is fuzziness parameter.

In this formulation, $|(x_i - c_j)|$ represents the distance metric used to assess the proximity of data points to cluster centers, typically the Euclidean distance.

The fuzzy membership is updated using Formula 2.

$$u_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|}\right)^{\frac{2}{m-1}}} \quad (2)$$

This formula recalibrates the degree of membership for each data point x_i based on its proximity to the cluster centers c_j. The greater the distance, the lower the membership degree, thus encouraging group assignments based on closeness rather than strict boundaries.

Meanwhile, the cluster center (c_j) is updated using Formula 3.

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^{m} x_i}{\sum_{i=1}^{N} u_{ij}^{m}} \quad (3)$$

This updates the position of each cluster center by taking a weighted average of all data points based on their membership degrees. The process iterates until either the change in the value of the objective function converges or a pre-defined iteration limit is reached.

The package used for FCM implementation is the e1071 package [30], which provides functions for running the FCM algorithm efficiently.

2.4 Optimization parameters of FCM using metaheuristics techniques

To obtain the most suitable number of clusters and fuzziness parameter for the FCM algorithm, this study utilizes metaheuristic optimization techniques. These methods are particularly effective for addressing complex and non-convex optimization problems, which are often encountered in the parameter tuning of clustering algorithms [24]. The primary objective of the optimization process is to maximize clustering quality, which is assessed using the Fuzzy Silhouette Index (FSI), a metric that evaluates the degree of separation between clusters while accounting for the fuzzy nature of membership.

In this study, three widely used metaheuristic algorithms were employed: Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and Differential Evolution (DE). The optimization was performed over two crucial parameters: the number of clusters k (ranging from 2 to 10) and the fuzziness coefficient m (ranging from 1.5 to 3). The FSI was used as the fitness function to evaluate the clustering result of each parameter combination.

The Genetic Algorithm (GA) simulates the natural evolutionary process to identify optimal solutions. It begins with a randomly initialized population of candidate solutions, represented as pairs of parameters (k, m). Through successive generations, these candidates evolve via selection, crossover, and mutation processes. In the selection phase, candidates with higher FSI scores are favored for reproduction, which ensures that superior solutions are likely to persist in the population. The crossover step allows selected candidates to exchange segments of their parameter configurations, creating new candidates. Meanwhile, mutation introduces random adjustments to certain candidates, ensuring genetic diversity. The optimization is conducted with a population size of 50 over up to 100 iterations, leveraging the GA package in R [31], [32].

Particle Swarm Optimization (PSO) is inspired by the social behavior of birds or fish in nature. Each particle represents a potential solution and moves through the search space by adjusting its trajectory based on its own best-known position and the best-known position among its neighbors [33]. The algorithm was run for a maximum of 100 iterations, with random initial positions for k and m. PSO was implemented using the pso package in R [34].

Differential Evolution (DE) operates by maintaining a population of candidate solutions and applying mutation and crossover operators to generate new solutions. Selection ensures that only the best candidates survive to the next generation. In this study, the DE algorithm was configured with the same parameter boundaries and iterated up to 100 generations using the DEoptim package in R [35].

In summary, each of these optimization techniques is employed to search for the best combination of k and m that empowers the FCI clustering process, yielding the highest FSI score and ensuring optimal fuzzy clustering performance for the school segmentation task. This comprehensive approach enhances the reliability of clustering results, which is crucial for developing effective educational marketing strategies.

2.5 Evaluation using Fuzzy Silhouette Index (FSI)

After obtaining FCM clustering results with optimal parameters from metaheuristic methods (GA, PSO, and DE), the next step is to evaluate the cluster quality using the Fuzzy Silhouette Index (FSI). The FSI is a refined measure compared to the conventional Silhouette Index because it explicitly accounts for the level of uncertainty inherent in cluster memberships [36]. This makes FSI particularly suitable for fuzzy clustering scenarios, where data points may belong to multiple clusters with varying degrees of membership.

The FSI is defined mathematically such as Formula 4.

$$FSI = \frac{\sum_{j=1}^{N}(\mu_{pj}-\mu_{qj})^{\alpha} s_j}{\sum_{j=1}^{N}(\mu_{pj}-\mu_{qj})^{\alpha}} \quad (4)$$

Where s_j is the Silhouette value for object j, μ_{pj} and μ_{qj} are the highest and second highest membership degrees, and α is the weighting coefficient. FSI is more accurate than the conventional Silhouette Index because it takes into account the level of probability in cluster membership. The package used to calculate FSI is the fclust package [37].

2.6 Membership Analysis and Marketing Strategy Formulation

After the clustering process, the next step involves analyzing the clustering results to formulate a more effective and adaptive marketing strategy. This process consists of three key components.

First, each cluster is assessed based on the average values of recency, frequency, and monetary (RFM) to determine its strategic marketing level. Clusters with high RFM scores represent schools with strong engagement and high recruitment potential, whereas those with lower RFM scores may require targeted interventions or relationship-building efforts.

Second, the membership degrees of schools across clusters are analyzed. Since FCM allows partial membership in multiple clusters, this analysis provides insight into how closely each school aligns with different cluster characteristics. Schools with dominant membership in one cluster indicate consistent interaction patterns, while those with distributed memberships may exhibit more diverse or evolving collaboration behaviors.

Third, customized marketing strategies are designed based on both cluster levels and membership distributions. Schools with high membership in high-value clusters can be prioritized with exclusive collaboration programs and loyalty initiatives. Meanwhile, schools with more distributed memberships may benefit from flexible and varied engagement programs. This fuzzy-based approach enables dynamic and data-driven educational marketing strategies tailored to the unique profiles of each school, thereby enhancing recruitment effectiveness and partnership opportunities [13].

Ultimately, this structured approach not only utilizes the clustering results effectively but also fosters stronger relationships between the university and schools, promoting a more collaborative and supportive recruitment environment. By recognizing the nuances in school engagement and adapting marketing initiatives accordingly, higher education institutions can optimize outreach efforts and improve overall student recruitment outcomes.

3. Results and Discussions

This section presents the results of the clustering process using the RFM approach combined with the Fuzzy C-Means algorithm. In this study, data analysis was conducted using R along with a combination of various packages to support model construction and evaluation.

To enhance clustering accuracy and ensure more reliable segmentation, a metaheuristic-based optimization was applied to determine the optimal initial parameters for the Fuzzy C-Means method. These findings are then analyzed through a discussion of their marketing implications based on the identified clusters.

3.1 Results

The results are presented sequentially to reflect the analysis workflow. The first step is to build an RFM model based on the collected school data, where each school is assigned a score to reflect its level of involvement in student recruitment activities. Next, three metaheuristic optimization techniques namely Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and Differential Evolution (DE) were evaluated to determine the optimal number of clusters and fuzziness parameters for the Fuzzy C-Means (FCM) algorithm. The performance of each optimization method was assessed using the Fuzzy Silhouette Index (FSI) to ensure clustering accuracy and reliability.

Furthermore, the characteristics of each RFM-based cluster were analyzed to reveal patterns of school behavior and institutional engagement. This analysis enabled the classification of schools into meaningful marketing segments such as Top Performer, Loyal Contributor, Emerging Star, and others. Fuzzy membership heatmaps were also created to visualize the degree of association of schools across multiple clusters. This visualization provides a better understanding of how schools exhibit overlapping characteristics, which supports more adaptive and targeted marketing strategies.

3.1.1 Recency, Frequency, and Monetary Score

The segmentation process begins with the calculation of recency (R), frequency (F), and monetary (M) scores for each of the 460 schools. These three dimensions quantify each school's engagement level with the university over the 2020–2024 period and are critical for understanding student recruitment dynamics. Recency gauges how recently a school has sent students, frequency indicates how often schools do so during the five-year period, and monetary reflects the total number of accepted students.

As shown in Table 2, a total of 460 RFM scores were calculated. For example, School S001 received a high Recency (4) and Frequency (5) score, showing regular and recent engagement, but a low Monetary (1) score, suggesting few students were ultimately admitted. In contrast, School S078, with moderate recency (3) and low frequency (2), scored high in monetary (4), indicating fewer interactions but a high yield of successful enrollments.

These RFM scores provide valuable input features for the clustering process, enabling targeted marketing strategies to be devised based on observed patterns of school engagement.

Table 2. RFM score each schools

ID School	Recency Score	Frequency Score	Monetary Score
S001	4	5	1
S078	3	2	4
…	…	…	…
S460	1	3	3

3.1.2 Performance of Metaheuristic Optimization Techniques

To identify the optimal number of clusters and the fuzziness parameter in the FCM algorithm such as formula 1, three metaheuristic optimization techniques were applied: Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and Differential Evolution (DE). Each method aimed to maximize the clustering quality using the Fuzzy Silhouette Index (FSI) such as formula 4 as the evaluation metric, where a higher FSI score indicates better-defined clusters. This step is crucial, as selecting appropriate parameters significantly influences the effectiveness of the clustering results.

Table 3. Performance of Metaheuristic Optimization Techniques

Metaheuristics Method	Best Parameter estimate	Number of Cluster	FSI Score
GA	k: 10; m: 2.42;	10	0.7554
PSO	k: 9; m: 2.83;	9	0.7640
DE	k: 8; m: 2.98;	8	0.8023

As shown in Table 3, the DE algorithm achieved the best performance with an FSI of 0.8023, producing eight clusters (k = 8) and a fuzziness parameter of m = 2.98. PSO with an FSI of 0.7640 and produced nine clusters, while GA produced the lowest FSI of 0.7554 with ten clusters. The results show that even though all three methods worked well for clustering, DE was chosen for further analysis because it created clearer and more distinct clusters.

These findings are consistent with [25], which demonstrated that DE often outperforms PSO and GA in clustering tasks, providing better-defined groups and more stable performance across varying datasets. The superior results achieved by DE can be attributed to its effective mechanism for exploring the search space, which leads to an optimal balance between exploration and exploitation in the parameter optimization process.

3.1.3 RFM-Based Cluster Characteristics

The results of school clustering using the FCM method optimized by DE are visualized in Figure 1. This graph shows the distribution of schools based on their interaction patterns with the institution, projected into two main dimensions. Each color represents a different cluster, with dots indicating the positions of schools within that cluster. The shaded areas reflect the spatial distribution of schools belonging to each group, highlighting the clear separation and lack of overlap between clusters. This indicates well-defined group characteristics and demonstrates the effectiveness of the clustering method employed.

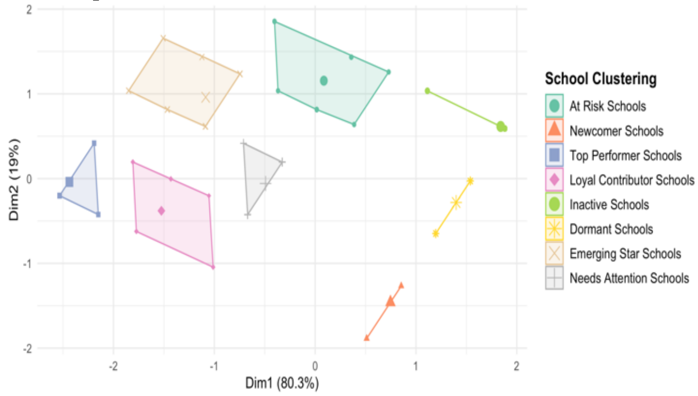

Figure 1. School clustering using FCM optimized by DE

Furthermore, the eight clusters were categorized into distinct marketing levels based on the average Recency, Frequency, and Monetary (RFM) scores. As shown in Table 4, Top Performer Schools exhibit the highest scores across all RFM dimensions, indicating consistent and significant student contributions. Loyal Contributor Schools also have high scores, particularly in recency and frequency, suggesting strong engagement with slightly lower student numbers than top performers. Emerging Star Schools demonstrate high frequency and monetary scores but moderate recency, reflecting growing potential that could be further developed.

Needs Attention Schools have medium scores across all RFM components, suggesting ongoing engagement that requires targeted strategies to improve. In contrast, Newcomer Schools display high recency but very low frequency and monetary scores, indicating new but limited involvement. At Risk Schools show low recency but medium frequency and monetary values, suggesting declining activity that warrants immediate intervention. Dormant Schools have moderate recency but very low scores in the other two dimensions, reflecting minimal recent involvement. Finally, Inactive Schools exhibit the lowest RFM scores, indicating nearly nonexistent interaction with the institution and a need for reactivation strategies.

The distribution of schools across clusters is illustrated in Figure 2. Dormant Schools represent the largest group, with 105 schools categorized under this cluster, highlighting that a significant proportion of schools have shown minimal engagement over the analyzed period. In contrast, At-Risk Schools comprise the smallest group, suggesting that only a few schools are currently on the verge of disengagement. Notably, Top Performer Schools, despite demonstrating the highest levels of consistent interaction, make up a relatively small portion of the total, highlighting that only a limited number of schools maintain strong, sustained partnerships with the institution. This analysis of cluster characteristics will inform the development of targeted marketing strategies tailored to enhance engagement and bolster recruitment efforts with various school segments.

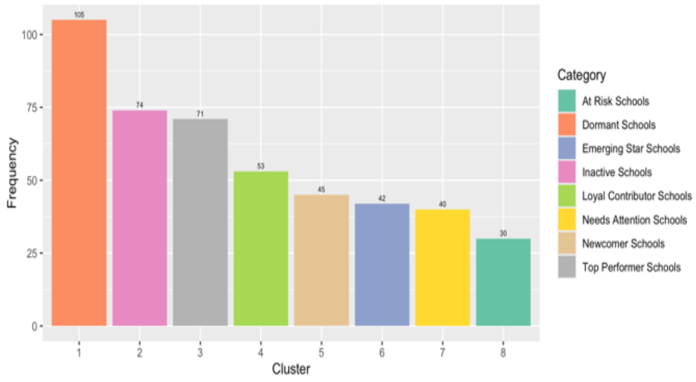

Figure 2. Number of schools in each cluster

Table 4. Marketing levels based on average of RFM each cluster

Cluste	$\bar{X}_R$	$\bar{X}_F$	$\bar{X}_M$	Marketing levels	Main Characteristics
1	1.53	3	3.23	At Risk Schools	R: low F: medium M: medium
2	4.31	1	1	Newcomer Schools	R: high F: very low M: very low
3	4.73	5	4.99	Top Performer Schools	R: very high F: very high M: very high
4	4.53	3.94	3.83	Loyal Contributor Schools	R: very high F: high M: high
5	1	1	1.11	Inactive School	R: very low F: very low M: very low
6	2.41	1	1	Dormant Schools	R: medium F: very low M: very low
7	2.57	4.26	4.12	Emerging Star Schools	R: medium F: high M: high
8	3.42	3	3.05	Needs Attention Schools	R: medium F: medium M: medium

3.1.4 Fuzzy Membership Distribution

With the fuzzy membership approach, school segmentation becomes more flexible and realistic, thereby enabling the design of adaptive marketing strategies. This approach contrasts with traditional hard clustering methods, in which each school is strictly assigned to a single cluster, limiting the ability to accurately reflect the complexities of school engagement. Table 5 presents the degree of membership for each school across the eight identified clusters produced by the Fuzzy C-Means (FCM) method. Unlike hard clustering, in which each school is assigned to only one cluster, the fuzzy approach allows a school to simultaneously belong to multiple clusters with varying degrees of membership.

For instance, School ID S460 shows a very high membership value of 0.962 in the Inactive Schools cluster, suggesting that this school is almost exclusively categorized as inactive. Conversely, School ID S002 has a more distributed membership profile, with scores such as 0.192 in the Top Performer cluster, 0.271 in Loyal Contributor, and 0.109 in Needs Attention. This indicates that the school exhibits characteristics that align with several clusters, reflecting a more complex behavioral pattern.

This flexible segmentation allows institutions to develop more targeted and dynamic marketing strategies. Schools with mixed memberships can receive multiple types of interventions tailored to their diverse characteristics. For example, School S002, being somewhat associated with both the Top Performer and Loyal Contributor clusters, could benefit from personalized engagement efforts that emphasize both academic excellence and loyalty-building initiatives.

Meanwhile, Figure 3 presents a heatmap visualization of the fuzzy membership distribution of each school across the identified clusters. The horizontal axis represents the eight cluster categories, while the vertical axis lists the school IDs (S001 to S460). Darker shades indicate higher degrees of membership, while lighter shades indicate lower levels.

Table 5. Fuzzy membership clustering

Cluster of School	School ID				
	S001	S002	S003	…	S460
At Risk	0.084	0.089	0.084	…	0.006
Newcomer	0.051	0.053	0.051	…	0.006
Top Performer	0.298	0.192	0.298	…	0.003
Loyal Contributor	0.207	0.271	0.207	…	0.004
Inactive	0.045	0.045	0.045	…	0.962
Dormant	0.050	0.052	0.050	…	0.010
Emerging Star	0.168	0.190	0.168	…	0.004
Needs Attention	0.097	0.109	0.097	…	0.005

From the heatmap, it is evident that certain clusters, such as the Inactive Schools, contain members with highly dominant cluster associations. For example, School ID S460 exhibits a strong membership of 0.962 in this cluster, reinforcing its categorization as inactive. On the other hand, clusters like Loyal Contributor Schools and Emerging Star Schools display more distributed membership patterns. This suggests that schools in these categories may share overlapping characteristics with other clusters. For instance, some schools categorized as Loyal Contributors also hold considerable membership in the Top Performer cluster, implying transitional attributes between the two.

Overall, the insights gathered from the fuzzy membership distribution not only enhance the understanding of school dynamics but also guide the institution in crafting tailored marketing strategies that foster deeper engagement and improve recruitment outcomes.

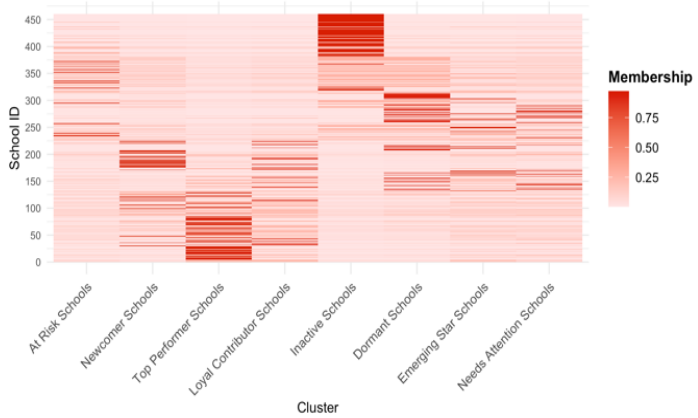

Figure 3. Heatmap of school fuzzy membership

3.2 Discussions

The results of the fuzzy clustering process utilizing the RFM approach and Fuzzy C-Means (FCM) method optimized by metaheuristic techniques, particularly Differential Evolution, successfully identified eight distinct school clusters. These clusters serve as a robust foundation for developing more segmented, dynamic, and data-driven educational marketing strategies. Unlike traditional hard clustering methods, the fuzzy approach provides greater flexibility in capturing the complexity of school behaviors by allowing institutions to have varying degrees of membership across multiple clusters. This finding aligns well with previous literature that emphasizes the importance of addressing ambiguity and uncertainty in segmentation, especially within marketing and customer analytics contexts [10], [13].

The resulting eight clusters, which are Top Performer, Loyal Contributor, Emerging Star, Dormant, Inactive, Newcomer, Needs Attention, and At Risk, reflect a diverse set of engagement patterns and relationship dynamics between schools and higher education institutions. Each cluster represents different levels of contribution and partnership potential, necessitating personalized and adaptive marketing approaches. For instance, Top Performer and Loyal Contributor schools are strongly and consistently engaged, making them ideal for long-term collaboration strategies such as loyalty rewards, priority partnerships, and exclusive academic dashboards. These findings support earlier studies that highlighted the need for institutions to cultivate strong relationships with loyal partners to maintain or boost student enrolment [38].

Conversely, the Dormant and Inactive clusters indicate minimal engagement, requiring reactivation strategies. Marketing initiatives focused on nostalgia, alumni involvement, and emotionally resonant storytelling may serve to reignite institutional affiliations among these schools. This insight corroborates the findings of [39], which assert that positive memory activation is effective in regaining the attention of previously disengaged customer segments.

Meanwhile, Emerging Star and Newcomer schools present valuable opportunities for growth, evidencing increasing engagement and a willingness to explore partnerships. Thus, strategies such as interactive campus experiences, hybrid visits, and technology-driven storytelling, such as chatbot assistance and video tours, can be effective in deepening relationships. Experience-based marketing strategies have been shown to significantly enhance brand engagement and conversion intention, especially among younger, digitally native audiences [40].

The Needs Attention and At Risk clusters highlight cases requiring targeted intervention strategies. These schools may exhibit signs of disengagement influenced by various institutional or external factors. For these segments, a data-informed approach involving performance reports, personalized consultations, and performance-based incentives may be pertinent. Research supports that adaptive relationship marketing strategies are particularly relevant for these clusters [41].

In contrast to the majority of existing research that applies clustering or segmentation primarily for static grouping or descriptive reporting, this study advances the field by integrating fuzzy membership analysis into actionable strategy formulation. This approach is significant: schools are not rigidly assigned to a single category but rather show varying degrees of association across multiple segments. This framework empowers institutions to implement multi-strategy marketing campaigns tailored to the overlapping traits of each school.

Table 6. Main marketing strategy and example of implementation

Cluster	Main marketing strategy	Implementation example
Top Performer School	Retention & loyalty reward	1. Exclusive MoU agreement 2. Invitation to the school to become a main partner 3. Special dashboard of performance and student achievement 4. Awarding “Best Partner School”
Loyal Contributor School	Strengthening relationships and collaborative experiences	1. Teacher Appreciation Week program 2. Workshop with teachers and lecturers 3. Exclusive content delivery: alumni e-magazine, scholarship catalog 4. Teacher referral program
Emerging Star School	Growth promoters & accelerated transition to excellent clusters	1. Video testimonials of students from the school 2. Hybrid visit to campus (virtual & physical) 3. Alumni-based “Explore Career” program
Dormant School	Relationship reactivation with nostalgia & alumni approach	1. Email or WhatsApp personal remarketing 2. Alumni reunion on campus 3. Special alumni pathway offer 4. “Once from here, now successful there” campaign with alumni data from schools
Inactive School	Reconstruct relationships from zero & rebuild brand awareness	1. Joint social activities (CSR) 2. Send physical brochures to schools with premium design 3. Invitation to campus open house 4. Video campaign "Why Should Your School Know Us Again?"
Newcomer School	Informative & attractive introduction to the institution	1. Welcome package (email, WhatsApp, e-catalog, campus tour video) 2. Invitation to attend the study program introduction event online 3. Chatbot system for questions and answers
Needs Attention School	Intervention based on analysis of previous performance	1. Exclusive consultation 2. Submit performance report for the last 3 years 3. Incentives for schools if they succeed in increasing student interest
At Risk School	Personal relationship survival strategy	1. Visits from campus leadership 2. Special early enrolment discount offer 3. Personalized video content: “Why students from your school are a good fit here”

For example, a school with strong membership in both the Loyal Contributor and Emerging Star clusters can benefit from simultaneous relationship-strengthening initiatives and innovative growth programs. This aligns with the principles of precision marketing in higher education that advocate for a flexible, data-driven approach to communication [13], [42].

This study is among the first to demonstrate how FCM clustering, enhanced with metaheuristic optimization (GA, PSO, DE), can be directly applied to educational data for the development of recruitment strategies. While several previous studies have applied Fuzzy C-Means in educational contexts [12], [14], [15], [16], [17], they have primarily focused on descriptive segmentation, without incorporating parameter optimization and translating the results into actionable marketing strategies. In contrast, this research not only optimizes the FCM process using metaheuristic techniques but also integrates the resulting fuzzy membership outputs into a dynamic, data-driven marketing framework aligned with various patterns of school engagement. This integration of optimized fuzzy clustering and RFM modelling addresses a notable gap in the current literature.

Overall, this study illustrates that adopting a fuzzy membership-based segmentation approach enhances the interpretability of school behaviors and facilitates the implementation of personalized, adaptive, and data-driven marketing strategies. This approach is poised to support the evolving practice of precision marketing in higher education, offering institutions a clearer path to improving recruitment outcomes. A summary of the main marketing strategies and their implementation examples for each cluster is presented in Table 6. Future research could explore the integration of qualitative feedback from schools to refine the segmentation and further validate the proposed marketing strategies.

4. Conclusions

This study aims to segment the location of new student admission promotions using the RFM approach and the Fuzzy C-Means method, with the expectation of producing a more adaptive and data-driven marketing strategy. The clustering results reveal eight distinct clusters: Top Performer, Loyal Contributor, Emerging Star, Dormant, Inactive, Newcomer, Needs Attention, and At Risk, which reflect the diverse relationships between schools and higher education institutions. The flexibility provided by the fuzzy approach allows for a nuanced understanding of school characteristics, as each institution can belong to multiple clusters with varying degrees of membership. Additionally, the optimization of Fuzzy C-Means parameters using Differential Evolution (DE) demonstrated improved clustering performance compared to traditional methods, thereby confirming that optimal parameter selection significantly enhances segmentation accuracy.

These findings suggest that marketing strategies based on optimized fuzzy segmentation are more effective than those derived from hard clustering approaches, as they encapsulate complex relational dynamics and lay the groundwork for more personalized and proactive marketing initiatives. This study contributes to the field of educational marketing by illustrating how data-driven segmentation can lead to targeted engagement strategies that foster deeper connections between institutions and schools.

However, this study has some limitations. The RFM variables selected for school segmentation can be further refined by incorporating additional factors such as demographic information, levels of participation in campus activities, and institutional feedback. Furthermore, this study primarily relied on historical data, leaving the actual impact of the proposed marketing strategies untested. This aspect requires further validation through experimental or longitudinal studies to accurately assess its effectiveness. It is important to note that the examples of marketing strategies provided in this study are context dependent and may vary significantly across different locations and conditions.

For future research, this framework could benefit from integrating other machine learning techniques, such as hybrid clustering, and exploring spatial effects to analyze how geographical location influences segmentation accuracy. Additionally, direct intervention studies could be employed to measure the impact of the proposed marketing strategies on applicant numbers from each school segment. Implementing real-time data analysis and developing AI-based systems for the automation of marketing strategy recommendations also represent promising avenues for supporting data-driven marketing within higher education institutions.

Ultimately, this study underscores the importance of adopting a flexible, fuzzy membership-based segmentation approach that not only enhances the interpretability of school behaviors but also empowers institutions to implement personalized, adaptive, and strategic marketing initiatives. By addressing the evolving landscape of higher education recruitment, this research paves the way for improved recruitment outcomes and stronger institutional partnerships.

Acknowledgements

This research was made possible through the support of STIEM Bongaya, which granted the research team access to its new student admission database. The authors sincerely thank the institution for its cooperation and openness in facilitating this study. The authors declare no conflict of interest. This research did not receive any specific grant from funding agencies in the public, commercial, or non-profit sectors.

References

[1] R. D. Dana, C. L. Rohmat, and A. R. Rinaldi, "Strategi Marketing Penerimaan Mahasiswa Baru Menggunakan Machine Learning dengan Teknik Clustering," Jurnal Informatika: Jurnal Pengembangan IT, vol. 4, no. 2–2, pp. 201–204, Dec. 2019, doi: 10.30591/JPIT.V4I2-2.1879.

[2] A. N. Alifah, R. P. Permata, I. M. W. A. Sanjaya, P. Arismawati, and W. A. Prastyabudi, "Data-Driven Insights into High School Market Segmentation Using Multidimensional Clustering Approaches," 2024 Beyond Technology Summit on Informatics International Conference, BTS-I2C 2024, pp. 47–52, 2024, doi: 10.1109/BTS-I2C63534.2024.10941845.

[3] Ernawati, S. S. K. Baharin, and F. Kasmin, "Target Market Determination for Information Distribution and Student Recruitment Using an Extended RFM Model with Spatial Analysis," Journal of Distribution Science, vol. 20, no. 6, pp. 1–10, 2022, doi: 10.15722/JDS.20.06.202206.1.

[4] W. A. Prastyabudi, A. N. Alifah, and A. Nurdin, "Segmenting the Higher Education Market: An Analysis of Admissions Data Using K-Means Clustering," Procedia Comput Sci, vol. 234, pp. 96–105, Jan. 2024, doi: 10.1016/J.PROCS.2024.02.156.

[5] A. Diana, A. Ariesta, A. Wibowo, A. B. Risaychi, and C. Author, "New Student Clusterization Based On New Student Admission Using Data Mining Method," Jurnal Pilar Nusa Mandiri, vol. 19, no. 1, pp. 1–10, Mar. 2023, doi: 10.33480/PILAR.V19I1.4089.

[6] D. Susilowati et al., "Segmentasi Lokasi Promosi Penerimaan Mahasiswa Baru Menggunakan Metode RFM dan K-Means Clustering," MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer, vol. 21, no. 2, pp. 275–282, Mar. 2022, doi: 10.30812/MATRIK.V21I2.1542.

[7] O. N. Akande, H. B. Akande, E. O. Asani, and B. T. Dautare, "Customer Segmentation through RFM Analysis and K-means Clustering: Leveraging Data-Driven Insights for Effective Marketing Strategy," International Conference on Science, Engineering and Business for Driving Sustainable Development Goals, SEB4SDG 2024, 2024, doi: 10.1109/SEB4SDG60871.2024.10630052.

[8] S. Safouan, K. El Moutaouakil, and A. M. Patriciu, "Fractional Derivative to Symmetrically Extend the Memory of Fuzzy C-Means," Symmetry 2024, Vol. 16, Page 1353, vol. 16, no. 10, p. 1353, Oct. 2024, doi: 10.3390/SYM16101353.

[9] A. Qtaish, M. Braik, D. Albashish, M. T. Alshammari, A. Alreshidi, and E. J. Alreshidi, "Optimization of K-means clustering method using hybrid capuchin search algorithm," Journal of Supercomputing, vol. 80, no. 2, pp. 1728–1787, Jan. 2024, doi: 10.1007/S11227-023-05540-5/METRICS.

[10] S. Sobantu and O. E. Isafiade, "Cascaded RFM-Based Fuzzy Clustering Model for Dynamic Customer Segmentation in the Retail Sector," Communications in Computer and Information Science, vol. 2326 CCIS, pp. 53–73, 2025, doi: 10.1007/978-3-031-78255-8_4.

[11] K. Zhou and S. Yang, "Effect of cluster size distribution on clustering: a comparative study of k-means and fuzzy c-means clustering," Pattern Analysis and Applications, vol. 23, no. 1, pp. 455–466, Feb. 2020, doi: 10.1007/S10044-019-00783-6/METRICS.

[12] S. Hidayat, R. Rismayati, M. Tajuddin, N. Luh, and P. Merawati, "Segmentation of university customers loyalty based on RFM analysis using fuzzy c-means clustering," Jurnal Teknologi dan Sistem Komputer, vol. 8, no. 2, pp. 133–139, Apr. 2020, doi: 10.14710/JTSISKOM.8.2.2020.133-139.

[13] M. Sivaguru and M. Punniyamoorthy, "Dynamic fuzzy clustering algorithm using cumulative cost curve with application to dynamic customer segmentation," Int J Bus Inf Syst, vol. 46, no. 1, pp. 73–106, 2024, doi: 10.1504/IJBIS.2024.138551.

[14] Y. Li, J. Gou, and Z. Fan, "Educational data mining for students' performance based on fuzzy C-means clustering," The Journal of Engineering, vol. 2019, no. 11, pp. 8245–8250, Nov. 2019, doi: 10.1049/JOE.2019.0938.

[15] I. Singh, S. Sabitha, T. Choudhury, A. Aggarwal, and B. K. Dewangan, "Mapping Student Performance With Employment Using Fuzzy C-Means," https://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/IJISMD.2020100103, vol. 11, no. 4, pp. 36–52, Jan. 1AD, doi: 10.4018/IJISMD.2020100103.

[16] B. A. Muhammad et al., "A Fuzzy C-means Algorithm to Detect Learning Styles in Online Learning Environment," Journal of Networking and Network Applications, vol. 4, no. 1, pp. 39–47, May 2024, doi: 10.33969/J-NANA.2024.040105.

[17] K. B. Utomo, P. Arbain, and Suminto, "Application of fuzzy C-means algorithm for basic school clustering in Samarinda city based on minimum educational service standard indicators," IOP Conf Ser Mater Sci Eng, vol. 885, no. 1, p. 012002, Jul. 2020, doi: 10.1088/1757-899X/885/1/012002.

[18] S. Deng, "Clustering with Fuzzy C-means and Common Challenges," J Phys Conf Ser, vol. 1453, no. 1, p. 012137, Jan. 2020, doi: 10.1088/1742-6596/1453/1/012137.

[19] Y. Chen, S. Zhou, X. Zhang, D. Li, and C. Fu, "Improved fuzzy c-means clustering by varying the fuzziness parameter," Pattern Recognit Lett, vol. 157, pp. 60–66, May 2022, doi: 10.1016/J.PATREC.2022.03.017.

[20] H. Zhang and S. L. Huang, "Improved fuzzy C-means clustering algorithm based on fuzzy particle swarm optimization for solving data clustering problems," Math Comput Simul, vol. 233, pp. 311–329, Jul. 2025, doi: 10.1016/J.MATCOM.2025.02.012.

[21] F. Belhabib, M. Benslimane, and K. El Moutaouakil, "Advanced Big Data Analytics: Integrating Fuzzy C-Means, Encoder-Decoder CNNs, and Genetic Algorithms for Efficient Clustering and Classification," Statistics, Optimization & Information Computing, vol. 13, no. 1, pp. 222–247, Jan. 2025, doi: 10.19139/SOIC-2310-5070-1978.

[22] S. Burada, B. E. Manjunathswamy, and M. Sunil Kumar, "Early detection of melanoma skin cancer: A hybrid approach using fuzzy C-means clustering and differential evolution-based convolutional neural network," Measurement: Sensors, vol. 33, p. 101168, Jun. 2024, doi: 10.1016/J.MEASEN.2024.101168.

[23] R. J. Kuo, C. C. Hsu, T. P. Q. Nguyen, and C. Y. Tsai, "Hybrid multi-objective metaheuristic and possibilistic intuitionistic fuzzy c-means algorithms for cluster analysis," Soft Computing 2023 28:2, vol. 28, no. 2, pp. 991–1008, Nov. 2023, doi: 10.1007/S00500-023-09367-3.

[24] K. Bhalla and A. Gosain, "Performance Analysis of Hybridized Fuzzy Clustering Algorithms Using Metaheuristic Algorithms," Lecture Notes in Networks and Systems, vol. 680 LNNS, pp. 445–461, 2023, doi: 10.1007/978-981-99-2602-2_34.

[25] A. Banerjee and I. Abu-Mahfouz, "A Novel Adaptive FCM with Cooperative Multi-Population Differential Evolution Optimization," Algorithms 2022, Vol. 15, Page 380, vol. 15, no. 10, p. 380, Oct. 2022, doi: 10.3390/A15100380.

[26] H. Wickham, R. François, L. Henry, K. Müller, and D. Vaughan, "dplyr: A Grammar of Data Manipulation," CRAN: Contributed Packages, Jan. 2014, doi: 10.32614/CRAN.PACKAGE.DPLYR.

[27] H. Wickham, "reshape2: Flexibly Reshape Data: A Reboot of the Reshape Package," CRAN: Contributed Packages, Apr. 2020, doi: 10.32614/CRAN.PACKAGE.RESHAPE2.

[28] R. C. Blattberg, B.-D. Kim, and S. A. Neslin, "RFM Analysis," Feb. 2024, doi: 10.1007/978-0-387-72579-6_12.

[29] J. C. Bezdek, R. Ehrlich, and W. Full, "FCM: The fuzzy c-means clustering algorithm," Comput Geosci, vol. 10, no. 2–3, pp. 191–203, Jan. 1984, doi: 10.1016/0098-3004(84)90020-7.

[30] D. Meyer, E. Dimitriadou, K. Hornik, A. Weingessel, and F. Leisch, "Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien [R package e1071 version 1.7-16]," CRAN: Contributed Packages, Sep. 2024, doi: 10.32614/CRAN.PACKAGE.E1071.

[31] B. Alhijawi and A. Awajan, "Genetic algorithms: theory, genetic operators, solutions, and applications," Evol Intell, vol. 17, no. 3, pp. 1245–1256, Jun. 2024, doi: 10.1007/S12065-023-00822-6/METRICS.

[32] L. Scrucca, "GA: A package for genetic algorithms in R," J Stat Softw, vol. 53, no. 4, pp. 1–37, 2024, doi: 10.18637/JSS.V053.I04.

[33] L. Abualigah et al., "Particle swarm optimization algorithm: review and applications," Metaheuristic Optimization Algorithms: Optimizers, Analysis, and Applications, pp. 1–14, Jan. 2024, doi: 10.1016/B978-0-443-13925-3.00019-4.

[34] C. Bendtsen, "pso: Particle Swarm Optimization," CRAN: Contributed Packages, Sep. 2022, doi: 10.32614/CRAN.PACKAGE.PSO.

[35] K. M. Mullen, D. Ardia, D. L. Gil, D. Windover, and J. Cline, "DEoptim: An R package for global optimization by differential evolution," J Stat Softw, vol. 40, no. 6, pp. 1–26, 2011, doi: 10.18637/JSS.V040.I06.

[36] R. J. G. B. Campello and E. R. Hruschka, "A fuzzy extension of the silhouette width criterion for cluster analysis," Fuzzy Sets Syst, vol. 157, no. 21, pp. 2858–2875, Nov. 2006, doi: 10.1016/J.FSS.2006.07.006.

[37] P. Giordani, M. B. Ferraro, and A. Serafini, "fclust: Fuzzy Clustering," CRAN: Contributed Packages, Nov. 2022, doi: 10.32614/CRAN.PACKAGE.FCLUST.

[38] A. Supriyanto et al., "Partnership Quality, Student's Satisfaction, and Loyalty: A Study at Higher Education Legal Entities in Indonesia," Open Education Studies, vol. 6, no. 1, Jan. 2024, doi: 10.1515/EDU-2024-0035/ASSET/GRAPHIC/J_EDU-2024-0035_FIG_002.JPG.

[39] P. Shukla et al., "The Role of Emotions in Consumer Brand Loyalty: A Neuromarketing Approach," International Journal of Tourism and Hospitality in Asia Pasific, vol. 7, no. 1, pp. 104–116, Feb. 2024, doi: 10.32535/IJTHAP.V7I1.2901.

[40] C. Prentice, X. Wang, and S. M. C. Loureiro, "The influence of brand experience and service quality on customer engagement," Journal of Retailing and Consumer Services, vol. 50, pp. 50–59, Sep. 2019, doi: 10.1016/J.JRETCONSER.2019.04.020.

[41] S. P. John and R. De Villiers, "Factors affecting the success of marketing in higher education: a relationship marketing perspective," Journal of Marketing for Higher Education, Jul. 2024, doi: 10.1080/08841241.2022.2116741.

[42] N. Behare et al., "Data-Driven Insights: The Power of Genetic Information in Precision Marketing," https://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/979-8-3693-2367-0.ch003, pp. 41–65, Jan. 1AD, doi: 10.4018/979-8-3693-2367-0.CH003.