e-Jurnal Apresiasi Ekonomi Volume 12, Nomor 3, September 2024: 576-584 ISSN Cetak : 2337-3997
ISSN Online : 2613-9774

DETERMINING THE MARKETING STRATEGY OF STIE MAHAPUTRA RIAU USING THE K-MEANS CLUSTERING ALGORITHM METHOD

Rahmadani Hidayat[1], Sarjon Defit[2], Menhard[3]

[13]Management, STIE Mahaputra Riau
email: *rahmadanihidayat83@gmail.com, email: menhard1967@gmail.com*
[2]Information Technology, Universitas Putra Indonesia YPTK Padang
email: *sarjond@yahoo.co.uk*

ABSTRACT

The difficulty of getting new prospective students requires STIE Mahaputra Riau to be able to design an effective and efficient marketing strategy. This study aims to determine a marketing strategy using the K-Means Clustering method. The K-Means Clustering algorithm method is to cluster data based on the attributes of student name, school of origin, area of origin and chosen study program, so that cluster data output is obtained that can be used in making marketing strategy decisions. The sample data used in this study are data from high school, vocational high school or equivalent students who are in the third grade in 2023, specifically for the province of Riau and its surroundings, totaling 750 data. The results of this study indicate that based on the total student data of 750 people, they are grouped into 3 clusters. Cluster 1 consists of 145 people from Rokan Hulu, Indragiri Hilir, Bengkalis, Kuantansingingi and West Sumatra Regencies. Cluster 2 consists of 344 people from Kampar and Indragiri Hulu Regencies. And cluster 3 as many as 261 people from Pelalawan, Siak and Rokan Hilir Regencies. It was also found in each cluster, the study program with the most interest was the S1 Management study program. So the marketing strategy implemented should pay attention to the area of origin and the study program chosen as the basis for implementing policies in accepting new prospective students.

Keywords : Data Mining, Marketing Strategy, Clustering, K-Means Method

INTRODUCTION

Currently, the use of information technology has penetrated all business sectors, starting from trade, industry, services and other economic sectors. The use of information technology produces very large amounts of data. However, very few of these business people are able to manage and utilize this data wisely. If processing so much data is done correctly and effectively, it will really help business people in exploring and finding important information that can be used as a reference when making policies, one of which is for the marketing sector. Marketing is the art and science of determining target markets and how to acquire, maintain and increase customers, by creating, distributing and communicating values (Hidayat et al., 2022).

Marketing is an important factor that must be considered in winning business competition, including in the education sector. Marketing has a very crucial role for private universities in particular, especially in facing increasingly tight competition and changes in the behavior of prospective students. Universities are required to be able to design effective and more targeted marketing strategies, therefore a proper method and plan are needed in deciding the marketing strategy policy to be taken. This can be done by managing and utilizing student data and information from various high schools, vocational schools or equivalent, which are then used as a basis for implementing marketing activities to find and obtain new prospective students at the university.

As a private university, the marketing strategy carried out by the STIE Mahaputra Riau Marketing Team is implemented through several stages, such as distributing scholarship invitation letters, conducting scholarship tests and distributing registration questionnaires to find out the interests of school students in the Riau Province and its surroundings. From this activity, data was obtained related to the biodata of the school students visited, while the data attributes used in this research were related to students' personal data, including school origin, regional origin and the department or study program they were interested in. The existing data must be managed and analyzed well, in order to produce information that is very useful in preparing marketing activities that are more targeted and more efficient in attracting prospective students..

One technology that can be used to process data is using data mining technology. Data mining is the process of discovering interesting knowledge from large amounts of data stored in databases, data warehouses, or other information stores (Han & Kamber, 2016). The data mining technology used in this research uses the K-Means clustering method to process and analyze existing data. K-Means clustering is a non-hierarchical data clustering, where with this method existing data will be grouped based on data that has similarities or similar data into the same cluster group, while data or objects that are different are grouped in different clusters with similar data characteristics (Santosa, 2007).

Data management using the clustering method is expected to produce data and provide useful information, so that the data can be used in determining appropriate and effective marketing strategies for the STIE Mahaptra Riau Marketing Team in searching for and recruiting prospective students.

Many studies have utilized data mining technology with the K-Means Clustering method to manage data from various industrial, trade and service sectors, including (Budiarti et al., 2006), In his research, he used clustering techniques with the Expectation Maximization (EM) algorithm in WEKA to be applied to the academic field of the Master of Information Technology Program at the University of Indonesia (MTI-UI) in processing data with gender and major type attributes. Based on the research conducted, it was found that An alternative route to accelerate student graduation is through a final project. (Bonnema & Van Der Waldt, 2008) using K-Means Clustering in the education sector to process and analyze data from 19 random schools to determine the preferences and information needs of a tertiary institution required by prospective students in South Africa. (Yanik & Elmorsy, 2013) grouping bank customer credit card transaction data using K-Means cluster to determine customer consumption levels such as food, entertainment, etc. (Tleis et al., 2017) grouping organic food consumers in Lebanon based on lifestyle and attitude constructs using k-means clustering.

Besides that (Wang et al., 2021) successfully utilized advanced K-means cluster center selection (FKM-ICS) feature solving the fault imbalance problem in vehicle fault diagnosis (VFT) of railway stations. Whereas (Baridam & Ali, 2013) using the K-means algorithm method with a pre-processor succeeded in grouping biological data mapped in low dimensions with good quality, compact and well separated. (Liu et al., 2021) in his research, he summarized the process of collective intelligence evolution into a DE solution process and used K-means clustering to identify individuals who were not beneficial to the evolution process in the early stages of intelligent evolution. (Jauhar et al., 2023) K-Means Clustering is used to group sellers based on the most sales and products offered in the e-commerce industry. (Yildirim et al., 2023) in his research in Turkey, he succeeded in grouping the similarities and differences between countries regarding national digital transformation (DT) and Industry 4.0 (I4.0) policies and according to the geographical, economic and political conditions of the country.

Various previous research references prove that there are many uses of the K-Means Clustering algorithm in grouping data for decision making. Therefore, in this research the author is interested in utilizing the K-Means Clustering algorithm to determine marketing strategies by processing and grouping school student data with the attributes of student name, regional origin and chosen study program as a basis for implementing policies in recruiting prospective new students.

RESEARCH METHODS

Clustering

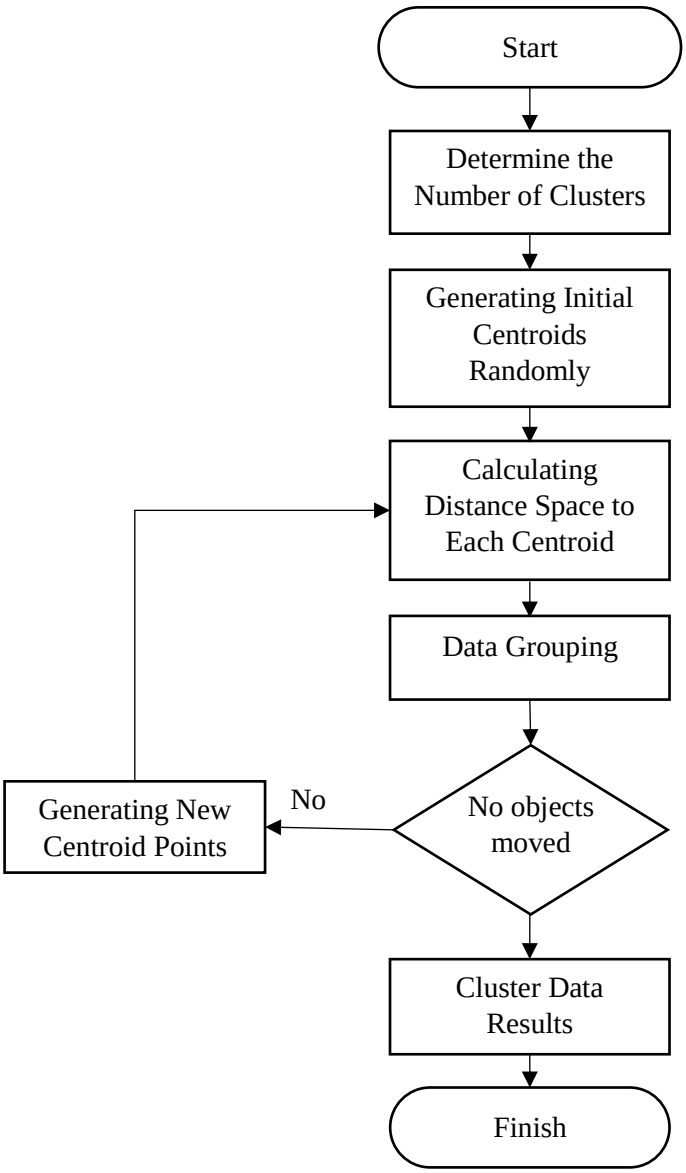
Clustering is the grouping of all forms of data that may be obtained from any field into categories that have similar characteristics for further analysis or to take action related to the field of study (Roohi, 2013). Clustering involves grouping records, observations, or concerns to form classes of similar objects. A cluster is a collection of records that are similar to each other and different from records in other clusters (Nugraha, 2014). There are two known clustering methods: hierarchical clustering and partitioning.

K-Means Algorithm

The K-means clustering algorithm is a non-hierarchical data clustering method that groups data into one or more clusters. Data that has the same characteristics is grouped into one cluster, and data that has different characteristics is grouped into another cluster, so that the variation of data in the cluster/group becomes less (Sulistyowati et al., 2018).

Stages of implementing the K-Means Clustering Algorithm according to (Nasari, Fina & Charles J. M., 2016) can be seen in figure 1.

ISSN Cetak : 2337-3997
ISSN Online : 2613-9774



Based on the image above, the steps for clustering using the K-Means method are as follows;

1) Determine the number of clusters
2) Determine the initial cluster center point randomly, which is then initialized with random numbers.
3) Map all data/objects to the nearest cluster. The proximity of two objects is determined by the distance between them. Likewise, the distance/proximity of data to a particular cluster is determined by the distance between the data and the cluster center. In this phase, we need to calculate the distance from all data to the center of each cluster. The distance between a data item and a particular cluster determines which cluster the data is in.
4) Then recalculate the cluster center using the current cluster membership. The center of a cluster is the average of all data/objects in a particular cluster. You can also use the cluster median (average) if you wish. Therefore, the mean (average) is not the only measure that can be used.
5) Reassign each object to use the new cluster center. If the cluster center does not change again, the clustering process is complete. Alternatively, return to step 3 until the cluster center no longer changes.

The research method used to determine the Marketing Strategy at STIE Mahaputra Riau is the application of the K-Means algorithm. The clustering process in the K-Means method has several stages as follows:

1. Data Collection

The sample data used in this research was obtained from the marketing department which carried out outreach activities for students in third grade in 2023 in high school, vocational or equivalent schools, especially for the Riau province and surrounding areas, totaling 750 data. This data contains all biographical data of school students, but in this research the author only uses a few data attributes to be used, such as student

name, school origin, region of origin, and choice of study program. The data for school students in 2023 can be seen in table 1 below.

Table 1. Example data of prospective students who are interested in continuing their education at STIE Mahaputra Riau in 2023

No	Student's name	Which school are you from	Place of Origin	Study Program
1	Aan Rianto	SMA Negeri 1 Inuman	Kuansing	S1 Management
2	Abdul Ghofur Azhari	SMA Negeri Cerenti	Kuansing	S1 Management
3	Abdul Hafiz	SMA 2 Batu Bersurat	Kampar	D3 Accounting
4	Abdul Mahdi	SMA Negeri 1 - Kampar kiri	Kampar	S1 Accounting
5	Acik Zeen	SMAN 1 Bangko Bagan Siapi-api	Rokan Hilir	D3 Accounting
6	Ade Setiawan	SMK YAPIM Tambusai Utara	Rokan Hilir	S1 Management
7	Adel Angela	SMA Negeri 2 Bunut	Pelalawan	S1 Management
8	Adil Mulya Candra	SMK Negri 1 Lubuk batu jaya	Indragiri Hulu	S1 Accounting
9	Adinda Yola Aprilia	SMK Perpajakan Riau	Pekanbaru	S1 Management
10	Adini Gulo	SMK Negeri 1 Bandar Sekijang	Pelalawan	S1 Accounting
11	Afasya Dwi Azuri	SMA Negeri 4 - Mandau	Bengkalis	S1 Accounting
12	Afdoli Fiqi	SMK Islam Inayah	Pekanbaru	S1 Accounting
13	Affan Muzakar Syah	SMK Islam Inayah - Ujung Batu	Rokan Hulu	S1 Management
14	Afriana Veronicha.S	SMKN 2 Rambah	Kampar	S1 Management
…				
…				
745	Zahra Anisa	SMA Negeri 1 - Minas	Siak	S1 Management
746	Zakaria Saragih	SMA Negeri 1 - kepenuhan	Rokan Hulu	S1 Management
747	Zakia Rahma	SMA Negeri 1 pangkalan kerinci	Pelalawan	S1 Management
748	Zakia Zahra Ramadhani	SMK N 1 TAMBANG	Kampar	S1 Management
749	Zovicka Julianda	SMK Baiturrahman - Kandis	Siak	S1 Management
750	Zulfikar Ali Butho	SMKN 1 Pangkalan Kerinci	Pelalawan	S1 Accounting

2. Data Preprocessing

The data preprocessing stage is carried out to prepare raw data before the analysis process is carried out using the k-means clustering method, by transforming nominal data such as area of origin, chosen study program and achievements which are then initialized into numerical form..

Data transformation

Data transformation is the process of changing data into data that can be processed by the algorithm that will be used, either in numerical form, classification, etc. (Nasari, Fina & Charles J. M., 2016).

The initialization process for the region of origin is carried out in the following stages:

1) In the regional data of origin, first group it by city or district, namely
   a. Kampar Regency consists of the areas of Batu Besurat, Tapung, Kampar Kiri, Bangkinang, Siak Hulu, Gunung Sahilan, and XIII Koto Kampar.
   b. Pekanbaru City which consists of schools in the city of Pekanbaru.

c. Pelalawan Regency which consists of Bandar Sekijang, Langgam, Bunut, Pangkalan Kerinci, Sorek, and Bandar Petalangan.
d. Siak Regency consists of Tualang, Minas, Kandis, Koto Gasib, Kerinci Kanan, Lubuk Dalam, and Bunga Raya.
e. Rokan Hilir Regency which consists of Bangko, Bagan Siapi-api, Pujud, and Rimba Melintang.
f. Rokan Hulu Regency which consists of Ujung Batu, Tambusai, Rambah, and Kecepatan.
g. Indragiri Hulu Regency which consists of Lubuk Batu Jaya, Peranap, Kelayang, Petai, Ukui, Lirik and Air Molek.
h. Bengkalis Regency which consists of the Duri, Kandis, Pinggir and Mandau areas
i. Kuantansingingi Regency which consists of Cerenti, Inuman, Singingi Hilir, and Kuantan Hilir.
j. West Sumatra Province which consists of the Pangkalan Koto Baru area.
k. North Sumatra Province which consists of the Mandailing Natal area.

2) The next step is to sort the data based on the largest number of students from each region of origin who are potential targets
3) After that, initials are given based on the order of origin, with the highest frequency given a value of 1, the second highest frequency with the number 2, and so on down to the lowest frequency. The results of this transformation can be seen in table 2.

Table 2. Transformation of Region of Origin data

Place of Origin	Frequency	Initials
Kampar Regency	236	1
Pekanbaru City	107	2
Pelalawan Regency	100	3
Siak Regency	99	4
Rokan Hilir Regency	63	5
Rokan Hulu Regency	56	6
Indragiri Hulu Regency	27	7
Bengkalis Regency	26	8
Kuantansingingi Regency	20	9
West Sumatra Province	16	10

In the same way, data initialization is also carried out to transform study program choice data in nominal form into data in numerical form, namely by sorting study programs with the highest frequency to the lowest frequency. So the results of the initialization of the program study are obtained which can be seen in table 3. Below;

Table 3. Initialization of Department Data

Study Program	Frequency	Initials
S1 Management	473	1
S1 Accounting	234	2
DIII Accounting	44	3

3. Cluster Analysis (K-Means Algorithm)

After the data is transformed into numbers, the data is grouped using the K-Means Clustering algorithm method. With this method, data is grouped based on data with the same characteristics which will be grouped into the same cluster, and another group for data with different characteristics. Steps for grouping data (clustering) using the K-Means method according to (Agusta, 2007) adalah sebagai berikut:

1) Determine the number of clusters,
2) Determine the initial centroid point value of each cluster.
3) Group the objects based on their closest distance to the cluster center. A data or object can be a member of a cluster, if it has the closest distance to the cluster centroid (Pramudyo, 2007). To calculate the distance of each data to each initial

cluster centroid, you can use Euclidean Distance theory with the following formula:

$$D(x,y)\sqrt{(X_1 - Y_1)^2 + (X_2 - Y_2)^2}$$

4) Determine the center of the new cluster again based on the average value of the cluster members. (Larose & Larose, 2014) in his book contains the determination of new centroid values using the following formula:

$$C = \frac{\Sigma m}{n}$$

Where:
C : data centroid
m : data members that belong to a particular centroid
n : the amount of data that belongs to a particular centroid

5) If data movement is still found among cluster members, then group the data or objects as in step three until there is no data movement from the cluster center.

RESULTS AND DISCUSSION

Results

Results of the Clustering Process in Iteration-1

1. The number of clusters determined in this research uses 2 clusters.
2. Determination of the initial centroid point.
   The initial centroid point (cluster center) is determined randomly based on the existing initial data. The initial cluster center selected can be seen in table 6 below;

Cluster	Name	Asal Sekolah	Place of Origin	Study Program
Cluster 1	Adini Gulo	SMKN 1 Bandar Seikijang	3	2
Cluster 2	Nur Uswatun Hasanah	SMKN 1 Kampar	1	1
Cluster 3	Trias Widya Lestari	SMAS Serirama YLPI	2	2

3. Calculate the distance space of all data to each cluster centroid using the Encludian Distance formula based on the centroid point of the first cluster C1 (3, 2), second cluster (1, 1) and third cluster (2, 2).
   Calculation of the distance of the first data to the first cluster centroid:

$$D(1,1)\sqrt{(9-3)^2 + (1-2)^2} = 6.08$$

From the calculation results it was found that the distance from the first data to the centroid of the first cluster was 6.08
Calculation of the distance from the first data to the second cluster centroid:

$$D(1,1)\sqrt{(9-1)^2 + (1-1)^2} = 8.00$$

From the calculation results it was found that the distance from the first data to the centroid of the second cluster was 8.00
Calculation of the distance from the first data to the third cluster centroid:

$$D(1,1)\sqrt{(9-2)^2 + (1-2)^2} = 7.07$$

From the calculation results it was found that the distance from the first data to the centroid of the third cluster was 7.07
Based on the results of calculating the distance between the first data and each cluster centroid, the lowest value obtained was 6.08, namely the distance between the first data and the centroid of the first culster, so it can be concluded that the closest distance to the first data is with cluster 1. Distance calculation results for the first 5 data and 5 data Finally, the three cluster centroids can be seen in table 7.

No	Student's Name	Place of Origin	Study Program	Cluster			Nearest Cluster
				C1	C2	C3	
1	Aan Rianto	9	1	6.08	8.00	7.07	C1
2	Abdul Ghofur Azhari	9	1	6.08	8.00	7.07	C1
3	Abdul Hafiz	1	3	2.24	2.00	1.41	C3
4	Abdul Mahdi	1	2	2.00	1.00	1.00	C2
5	Acik Zeen	5	3	2.24	4.47	3.16	C1
…							
…							
746	Zakaria Saragih	6	3	3.16	5.39	4.12	C1
747	Zakia Rahma	3	1	1.00	2.00	1.41	C1
748	Zakia Zahra	1	1	2.24	0.00	1.41	C2

	Ramadhani						
749	Zovicka Julianda	4	1	1.41	3.00	2.24	C1
750	Zulfikar Ali Butho	3	3	1.00	2.83	1.41	C1

1. The next step is to recalculate the new cluster centroid by calculating the average value of the data in the same cluster group for each cluster.
2. The results of calculating the new cluster centroid in the second iteration, with the first new cluster centroid values C1 (5.13, 1.56), C2 (1.23, 1.24) and C3 (1.53, 2.64).
3. Based on total data of 750 students with attributes of region of origin and study program, and using 3 cluster centroid points. The iterative process of calculating the distance of each data to the cluster centroid point is carried out until the 11th iteration, where during the calculation process from iteration 10 to iteration 11 there is no movement of the cluster centroid point and no data moves between clusters, so at this stage the iteration process is considered finished. From the results of these calculations, it can be seen that the data included in cluster 1 is 145 students, cluster 2 is 344 students, and cluster 3 is 261 students with the cluster centroid point in the 10th iteration, namely with values C1 (7.40, 1.59), C2 (1.31, 1.60).) and C3 (3.86, 1.54), as can be seen in table 8.

Table 8. Clustering Analysis Results

Cluster 1 Results	Cluster 2 Results	Cluster 3 Results
Cluster 1 consists of 145 people, who come from the region	Cluster 2 numbered 344 people, who came from the regions	Cluster 3 numbered 261 people, who came from the regions
Rokan Hulu Regency = 56 people	Kampar Regency = 237 people	Pelalawan Regency = 99 people
Indragiri Hulu Regency = 27 people	Indragiri Hulu Regency = 107 people	Siak Regency = 99 people
Bengkalis Regency = 26 people	And interested in choosing a study program	Rokan Hilir Regency = 63 people
Kuantansingingi Regency = 20 people	S1 Management = 195 people	And interested in choosing a study program
Prov. West Sumatra = 16 people	S1 Accounting = 93 people	S1 Management = 161 people
	D3 Accounting = 56 people	S1 Accounting = 58 people
And interested in choosing a study program		D3 Accounting = 42 people
S1 Management = 84 people		
S1 Accounting = 36 people		
D3 Accounting = 25 people		

Based on the table above, it can be seen that from the student data of several schools totaling 750 people, it can be grouped into 3 clusters. The results of this cluster provide an overview for the STIE Mahaputra Marketing Team in carrying out promotional activity strategies based on the distribution of district areas that are the main potential in recruiting prospective students. The area with the most interest comes from cluster 2 from Kampar and Indragiri Hulu Regencies, then with the second most interest, namely cluster 3 from Pelalawan, Siak and Rokan Hilir Regencies, and cluster 1 from Rokan Hulu, Indragiri Hilir, Bengkalis, Kuantansingingi and West Sumatra Province. The results of this analysis also show that the study program that is the most popular choice is the S1 Management study program. Therefore, from the results of this clusterization, STIE Mahaputra can form a promotion team and design marketing content based on the area of origin of prospective students and the most popular study programs, and can also utilize students from areas that are potential in recruiting prospective students.

Discussion

Based on the results of the clustering analysis in table 8 above, it can be seen that in cluster 1 the student characteristics are dominated by students who come from the Rokan Hulu district and Indragiri Hulu district. From the results of cluster 1, it can also be seen that the Bachelor of Management study program is the most dominant choice. So it can be concluded that in cluster 1, students generally come from the Rokan Hulu and

Indragiri Hulu districts who are interested in choosing the Bachelor of Management and Bachelor of Accounting study programs.

In the analysis results for cluster 2 with a total of 344 students, the students were dominated by students from the Kampar district area, with the largest choice of study programs dominated by Bachelor of Management and Bachelor of Accounting study programs. So it can be concluded that those in cluster 2 generally come from the Kampar district area and are interested in choosing the Bachelor of Management and Bachelor of Accounting study programs.

Meanwhile, the results of the analysis in cluster 3 with a total of 261 students, were dominated by students from two regions, namely Siak district and Pelalawan district, with the dominant study program choice being Bachelor of Management. So it can be concluded that students in cluster 3 generally come from the Siaj district and Pelalawan district, with a dominant interest in choosing the Bachelor of Management study program.

The results of this study are in line with research (Mahdiraji et al., 2022) uses the K-Means Cluster method to process customer demographic data and financial transactions which are then categorized into very loyal, loyal, highly involved, low involved and lost customers, thereby helping the bank to understand customer behavior patterns. The next researcher, namely (Güçdemir & Selim, 2015), carried out a business customer grouping approach using the hierarchical k-means clustering algorithm to determine market segments for original equipment manufacturers (OEM), by grouping data into five segment criteria starting from best, valuable, average, potentially valuable, and potentially not valuable. The research results state that the application of the k-means clustering algorithm can be used effectively and efficiently in determining business customer segmentation.

Marketing strategy of STIE Mahaputra Riau

Marketing activities carried out by STIE Mahaputra Riau are in the form of promotional activities. Promotion is an effort made to provide information, persuade and influence customers to accept and choose the products or services offered (Hidayat et al., 2023). Referring to the results of the clustering analysis using the k-means clustering algorithm method above, we can design several forms of promotional activities that can be carried out by STIE Mahaputra Riau in recruiting prospective new students, including;

1. Form a promotion team involving STIE Mahaputra Riau students who come from more dominant areas based on existing clusters. With this strategy, the marketing department can group students involved in promotional activities based on the student's regional origin which is adjusted to the origin of the school area that is the target of the promotion according to each cluster, so that by carrying out this strategy the marketing department can more easily carry out socialization and approaches. approach to school students and is a great opportunity to recruit prospective students.
2. Another strategy that the marketing team can use is carrying out promotional activities based on prospective students' interest in the choice of study program. This strategy is carried out by training and equipping students or marketing teams involved in promotional activities to have insight and knowledge according to the study program which is the dominant choice for students at the school so that the delivery of information is more targeted and can attract the interest of the school's students to continued his education at STIE Mahaputra Riau.

CONCLUSION

Based on the results of this research, it can be concluded that in carrying out its marketing activities, STIE Mahaputra Riau can carry out several promotional strategies, including; Form a Promotion Team involving STIE Mahaputra Riau students who come from areas based on existing clusters, as well as grouping the students involved based on the student's regional origin which is adjusted to the area of origin of the school that is the target of the promotion according to each cluster. Group students involved in promotional activities based on the student's study program, then provide students and other marketing teams with insight and knowledge according to the study program that prospective students are interested in based on the grouping in each existing cluster.

BIBLIOGRAPHY

Agusta, Y. (2007). *K-Means – Penerapan, Permasalahan dan Metode Terkait*. *3*(Pebruari), 47–60.

Baridam, B., & Ali, M. M. (2013). *An investigation of K -means clustering to high and multi-dimensional biological data*. https://doi.org/10.1108/K-02-2013-0028

Bonnema, J., & Van Der Waldt, D. L. R. (2008). Information and source preferences of a student market in higher education. *International Journal of Educational Management*, *22*(4), 314–327. https://doi.org/10.1108/09513540810875653

Budiarti, A., Sucahyo, Y. G., & Ruldeviyani, Y. (2006). *Studi karakteristik kelulusan peserta didik dengan teknik*. 199–207.

Güçdemir, H., & Selim, H. (2015). Integrating multi-criteria decision making and clustering for business customer segmentation. *Industrial Management & Data Systems*, *115*(6), 1022–1040. https://doi.org/https://doi.org/10.1108/IMDS-01-2015-0027

Han, J., & Kamber, M. (2016). *Data Mining: Concepts and Techniques*.

Hidayat, R., Menhard, M., Charli, C. O., & Masnum, A. (2023). *Determination Purchase Intention and Purchase Decision : Brand Image and Promotion Analysis (Marketing Management Literature Review)*. *4*(1), 152–165.

Hidayat, R., Piliang, A., Safrizal, S., & Sugianto, S. (2022). Pengaruh Harga Dan Kualitas Pelayanan Terhadap Kepuasan Konsumen Pada Hotel Cititel Sisingamangaraja Pekanbaru. *Jurnal Ilmiah Multidisiplin*, *1*(05), 131–136. https://doi.org/10.56127/jukim.v1i05.523

Jauhar, S. K., Chakma, B. R., Kamble, S. S., & Belhadi, A. (2023). Digital transformation technologies to analyze product returns in the e-commerce industry. *Journal of Enterprise Information Management*. https://doi.org/https://doi.org/10.1108/JEIM-09-2022-0315

Larose, D. T., & Larose, C. D. (2014). *Discovering Knowledge in Data* (2nd ed). Wiley. https://www.perlego.com/book/996125/discovering-knowledge-in-data-an-introduction-to-data-mining-pdf

Liu, J., Liang, B., & Ji, W. (2021). An anomaly detection approach based on hybrid differential evolution and K-means clustering in crowd intelligence. *International Journal of Crowd Science*, *5*(2), 129–142. https://doi.org/10.1108/IJCS-07-2020-0013

Mahdiraji, H. A., Tavana, M., Mahdiani, P., & Kamardi, A. A. A. (2022). A multi-attribute data mining model for rule extraction and service operations benchmarking. *Benchmarking: An International Journal*, *29*(2), 456–495. https://doi.org/https://doi.org/10.1108/BIJ-03-2021-0127

Nasari, Fina, A., & Charles J. M., S. (2016). Penerapan Algoritma K-Means Clustering Untuk Pengelompokkan Penyebaran Diare Di Kabupaten Langkat. *Cogito Smart Journal*, *2*(2), 108–119. https://doi.org/10.31154/cogito.v2i2.19.108-119

Nugraha, J. A. M. (2014). *Data Mining dengan Metode Clustering untuk Pengolahan Informasi Persediaan Obat pada Puskesmas Pandanaran Semarang*. https://api.semanticscholar.org/CorpusID:60828063

Pramudyo, A. S. (2007). *Case Based Reasoning untuk klasifikasi mahasiswa baru berdasarkan prediksi indeks prestasi pada semester I:: Studi kasus Program Studi Teknik Informatika Universitas Bina Darma Palembang*. Universitas Gadjah Mada.

Roohi, F. (2013). Artificial neural network approach to clustering. *The International Journal of Ngineering and Science*, *2*(3), 33–38.

Santosa, B. (2007). *Data Mining: Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Graha Ilmu.

Sulistyowati, Ketherin, B. E., Arifiyanti, A. A., & Sodik, A. (2018). *Analisa Segmentasi Konsumen Menggunakan Algoritma K-Means Clustering Jurusan Sistem Informasi , Institut Teknologi Adhi Tama Surabaya*.

Tleis, M., Callieris, R., & Roma, R. (2017). Segmenting the organic food market in Lebanon: an application of k-means cluster analysis. *British Food Journal*, *119*(7), 1423–1441. https://doi.org/https://doi.org/10.1108/BFJ-08-2016-0354

Wang, B., Wang, G., Wang, Y., & Lou, Z. (2021). *A K-means clustering method with feature learning for unbalanced vehicle fault diagnosis*. *2018*. https://doi.org/10.1108/SRT-01-2021-0003

Yanik, S., & Elmorsy, A. (2013). *SOM approach for clustering customers using credit card transactions*. https://doi.org/10.1108/IJICC-11-2018-0157

Yildirim, N., Gultekin, D., Hürses, C., & Akman, A. M. (2023). Exploring national digital transformation and Industry 4.0 policies through text mining: a comparative analysis including the Turkish case. *Journal of Science and Technology Policy Management*. https://doi.org/https://doi.org/10.1108/JSTPM-07-2022-0107