

Target Marketing Strategies Using Computer Based analysis in Profiling Potential School

A P Purfini, R Yunanto

Department of Computerized Accounting, Universitas Komputer Indonesia, Jl.Dipatikur 112-116 Bandung, Indonesia

E-mail : apriani.puti.purfini@email.unikom.ac.id

Abstract. The purpose of this research is to obtain information about schools that have good prospects for becoming university marketing targets. So schools can work directly with universities that target them. University can directly target high-achieving students from schools that have good reputations too. This study used data mining clustering methods to analyze the data of prospective new students and the data of graduating students. Data input for the data mining process used segmentation that considers retention, frequency, and average grade point average (GPA), this was needed to determine school characteristics. The results of this study are school profiles to identify a list of potential schools that will be targeted by university marketing.

1. Introduction

Universities need to manage educational organizations such as businesses so that they must be supported by the right strategies and tactics so that they can create a sustainable competitive advantage for the university. Identifying target markets is a key step in the process of attracting prospective students [1]. [2] emphasizing that market opportunity analysis is very useful providing a foundation for the interests of planning and marketing strategies. Managers are considered necessary to evaluate the environmental situation, the place where they devote their marketing activities and estimate the environmental influences that may arise in the future for consumers. Marketing to determine targets requires to focus on one or several selected market segments, and separate marketing programs for each segment [3]. Segmentation is performed on an unordered customer dataset and is the process of separating markets into groups of potential customers with similar needs and or characteristics. Market segmentation is an essential and popular marketing tool that can offer firms insights for growth (ie new product ideas) and efficiency (ie better marketing communication, better focus on the right audience [4]. The environment and the flow of change must be carefully monitored and evaluated, potential opportunities (opportunities) and problems that may arise must be identified before an organization develops a marketing plan.

Previous research discussed marketing strategies in educational institutions using several approaches. This research makes a method to find market segments among prospective students, and researchers build a conceptual model based on demographics [3]. This paper discusses prospective student market segments based on social media and examines the impact of social media on high school choice [5]. [6] Researching university preferences uses a form of conjoined analysis, known as adaptive conjoined analysis (ACA), to investigate a number of attributes for secondary school graduates in Australia. [7] The purpose of this study consisted of identifying factors that influence university choice



and academic programs by enrollees. [8] explores the concept of adapting the business relationship marketing framework to the challenges of college student retention.

This study aims to obtain market opportunity information in the form of information on which schools have the prospect of being the main target of marketing by analyzing and processing data on prospective new students and graduate data. Analysis and processing of data in this study used data mining techniques. The k-means algorithm is used to identify groups from schools based on retention, frequency, and average grade point average (GPA) values as input data. This is needed to know the characteristics of schools in various cities.

2. Methods

The flow in this study was adjusted to the stages in the Cross-Industry Standard for Data Mining (CRISP-DM) method. There were three stages of research; first was the stage of data collection (Library Study, Observation, Interview), stages of business understanding and stages of understanding data, Modeling and Evaluation Phase. This study used Binning to eliminate noise after data integration. Binning was used to reducing variations in data. The binning result dataset was then input data for data mining. This study used the K-Means data mining clustering method. The k-means clustering algorithm was the most commonly used [9] because of its simplicity [10]. The output of K-Means depends on the selected center values on clustering [11]. Consumer clustering was a process of identifying consumers to smaller and specific groups to make it easier to understand [12]. Clusters must be heterogeneous inside and desired among them [13].

3. Results and Discussion

Details of Business Needs and Data Source: Addition and increase in product capacity, Reduction of company operating costs, Increased marketing effectiveness, and profits. This research focuses on increasing marketing effectiveness and will subsequently have an impact on reducing the company's operating costs related to marketing activities. Business needs at the university include Profiling the school must record all business processes regarding the information on students who enroll as students and information about the acquisition of Grade Point Average (GPA) and year of graduation. Can see the names of schools from students who are students of the University. Can see which school data the frequency of students being University students. Can see which school data students have academic values, namely Grade Point Average (GPA) with predetermined criteria.

Transactional data used is structured data; data is taken from several databases that store the transaction. Based on the business needs of the University, the data domain consists of data on new student admissions and graduate data. Details of the data include New student admission data: level, study program, status, year of entry, registration number, name, personal identity number, place of birth, date of birth, sex, religion, address, origin of school, final score, father's job, mother's job, income, address parents, city, zip code ; data of graduates: level, study program, name, place of birth, date of birth, year of entry, year of graduation, grade point average (GPA). This pre-process stage includes data cleaning (data cleaning), data integration (data integration), task-relevant data, namely selecting data that has relevant attributes. By conducting data, the selection will help the stages of the data mining process in finding useful data patterns, which is a transformation process.

In cleaning the data, incomplete data on the contents were not included in the study. And complete data contents, selected several attributes so that the data to be processed is relevant to the needs. Thus it will improve performance in the mining process. Task relevant data is to select data that has relevant attributes. By conducting data, the selection will help the stages of the data mining process in finding useful data patterns. Therefore, not all data attributes will be used, only the attributes considered by the researcher are useful, and the distribution is not too random. Attributes that are relevant to defined business needs.

The next step is data integration. Data integration is the process of converting or merging data into a format suitable for processing in data mining. Often the data that will be used in the data mining process has a format that cannot be used immediately. Therefore, the format needs to be changed. In this study,

data with nominal data types are initialized into numbers so that they can be processed using the K-means Clustering algorithm. This stage is in line with the relevant data task process. Data transformation is required for data in order to obtain a dataset that fits your needs. This study considers school segmentation based on retention value, frequency, and average acquisition of grade point average (GPA). Retention to see the last time school students become students in the University, in this study the value of retention is the last year of entry for school students to become students at the University. Frequency to see the number of students who become students and graduate in the Vo University. The average grade point average (GPA) is obtained to see the average grade point average (GPA) of school students who are students and graduate at the University.

After data integration, it turns out there is noise data. Noise data will reduce the performance of the data mining algorithm. This research uses Binning to eliminate noise. Binning is used to reducing variations in data. The binning result dataset is then input data for data mining. Implementation and Evaluation in segmentation. From the dataset generated in the previous process, the data mining process will then be carried out. This study uses the IBM SPSS Modeler tool to do data mining. The process of data mining can be seen in Figure 1.

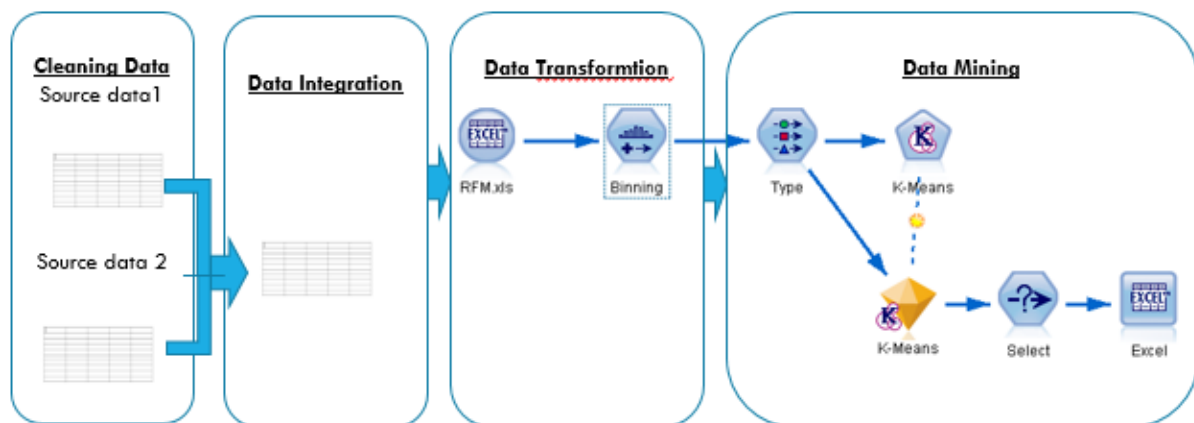


Figure 1. Data Integration

This research for data mining uses K-Means Clustering to the segment. K is the number of clusters that will be formed. To determine the appropriate K value, several trials were conducted from K = 3 to K = 7, then the best cluster quality was sought. Good cluster quality is produced with a silhouette approaching 1. Based on the experimental results, the value of K = 4 produces the best silhouette value compared to the other K values, shown in Figure 2.

Model Summary

Algorithm	K-Means
Inputs	3
Clusters	4

Cluster Quality

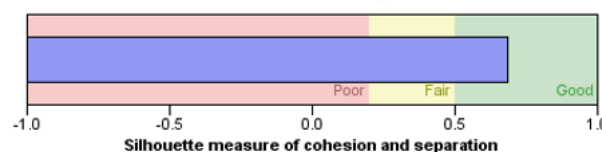


Figure 2. K-Means Result

Based on the results obtained in the data mining process, proceed to the next stage of the knowledge presentation. Knowledge presentation that is presenting patterns or information found that match facts or hypotheses that exist before. This pattern will later be useful as the output of this research. The output will answer the problem that has been formulated at the beginning because in this process, a formulation of the decision or action is taken from the results obtained. The output of the data mining process in this study is in the form of cluster data that identifies the segments of each school. This can be seen in Figure 3.

Cluster	cluster-1	cluster-3	cluster-2	cluster-4
Label				
Description				
Size	70.7% (29)	22.0% (9)	4.9% (2)	2.4% (1)
Inputs	F_BIN 1 (100.0%)	F_BIN 1 (100.0%)	F_BIN 10 (100.0%)	F_BIN 10 (100.0%)
	R_BIN 4 (100.0%)	R_BIN 1 (100.0%)	R_BIN 1 (100.0%)	R_BIN 4 (100.0%)
	M_BIN 4 (45.5%)	M_BIN 1 (33.3%)	M_BIN .	M_BIN 3 (100.0%)

Figure 3. Cluster data

Cluster 1, 70.7% of schools in this cluster with characteristics of low frequency, moderate retention, and moderate grade point average (GPA) values. Low frequency identifies schools that have rarely become students at the university, while Recency is identifying if students in the school have not been students in University for a long time and the score of the category Grade Point Average (GPA) is identifying that the school student has academic achievement potential well. If seen from the criteria, it can be identified that the school is less potential to be targeted for marketing.

Cluster 2, 22% of schools in this cluster with characteristics of low frequency, low retention, and GPA values. Low frequency identifies schools that have rarely become school students, while low recency identifies that students in the school have not been students at the university for a long time and the low category GPA scores identify that the school student has poor academic potential. If seen from the criteria, it can be identified that the school is less potential to be targeted for marketing. Cluster 3, 4.9% of schools in this cluster with characteristics of high frequency, low retention, and low GPA values. High-frequency schools often identify students as students at the university, while low recency identifies that students in these schools have not been students at the university for a long time and the low category GPA scores identify that the school student has poor academic potential. When viewed from the criteria, it can be identified that the school has the potential to be targeted for marketing.

Cluster 4, 2.4% of schools in this cluster with characteristics of high frequency, moderate retention, and low GPA values. High-frequency schools often identify those students as students at the university, while recency is identifying if students in the school have not been students at the university for a long time and the acquisition of low category GPAs identifies that the school student has poor academic potential. If seen from the criteria, it can be identified that the school is very potential to be targeted for marketing. The thinking framework of this study is under the explanation that marketing to decide targets requires to focus on one or several specific market segments, and separate marketing programs for each segment [3]. The method used in this study is different from earlier studies [3-6] without considering retention, frequency, and grade point average (GPA) values.

4. Conclusion

The method used in this study can be used to find potential new students based on school profiles. With the list of potential school, candidates can reduce the promotional costs that must be issued by UNIKOM. It is expected that with more specific targets, new student admissions at the university can be increased. The method used in this study can be used for further researchers to bundle together several study programs to be socialized in the target school.

Acknowledgement

This research was supported by Universitas Komputer Indonesia, Indonesia.

References

- [1] P W H. a. E. G. Miller, An approach for identifying benefit segments among prospective college students, *Journal of marketing for higher education*, **3**(1), pp. 49-65, 1990.
- [2] D. L. Boone L E, Contemporary Marketing Plus, Texas: Dryden Press, 1995.
- [3] D. M. Lewison and J. M. Hawes, 2007. Student target marketing strategies for universities," *Journal of college admission*, pp. 14-19.
- [4] Vidden, Chad, M. Vriens and a. S. Chen., 2016. Comparing clustering methods for market segmentation: A simulation study," *Applied Marketing Analytics*, **2**(3), pp. 225-238.
- [5] E. Constantinides and M. C. Zinck Stagno, 2011. "Potential of the social media as instruments of higher education marketing: a segmentation study," *Journal of marketing for higher education*, **21**(1), pp. 7-24.
- [6] G. N. Soutar and J. P. Turner, 2002. "Students' preferences for university: a conjoint analysis," *International Journal of Educational Management*, **16**(1), pp. 40-45.
- [7] A. M. Kalimullina and S. G. Dobrotvorskaya, 2016. "Higher Education Marketing Strategies Based on Factors Impacting the Enrollees' Choice of a University and an Academic Program," *International Journal Of Environmental & Science Education*, **11**(13), pp. 6025-6040.
- [8] R Ackerman and J. Schibrowsky, 2008. A Business Marketing Strategy Applied To Student Retention: A Higher Education Initiative, *J College Student Retention*, **9**(3), pp. 307-336.
- [9] R. C. D. a. A. K. Jain. 1988. Algorithms for Clustering Data, Prentice Hall.
- [10] A. Ng 2012. Clustering with the K-Means Algorithm, Machine Learning.
- [11] M. A. Syakur, B. K. Khotimah, E. M. S. Rochman and B. D. Satoto, 2018. Integration K-Means Clustering Method and Elbow Method For Identification of The Best Customer Profile Cluster. *IOP Conference Series: Materials Science and Engineering*, **336**(1), p. 012017.
- [12] A. Mohammad and M. Shahriar. 2010. Mining the banking customer behavior using clustering and association rules methods," *International Journal of Industrial Engineering & Production Research*, **21**, pp. 239-245.
- [13] K. Anil, 2009. Data clustering : 50 years beyond k-means. *International Journal of pattern recognition letters*.

Reproduced with permission of copyright owner. Further reproduction prohibited without permission.