



Customer Relationship Management in Retail Using Double Association Rule

Rahmat Abadi Suharjo¹, Antoni Wibowo²

¹Computer Science Department, BINUS Graduate Program – Master of Computer Science, Bina Nusantara University, Jakarta, Indonesia 11480, rahmat.suharjo@binus.ac.id

²Computer Science Department, BINUS Graduate Program – Master of Computer Science, Bina Nusantara University, Jakarta, Indonesia 11480, anwibowo@binus.edu

ABSTRACT

Retail is one of the industries that has very high competition. Many retail industries have increased so that there is high competition among retailers. The competition in the retail industry is centered on the struggle for consumers. Consumers are one of the main pillars for the survival of a company. Customer relationship management is a strategy to maintain customer loyalty. To maintain consumer loyalty by knowing consumer shopping habits. Market basket analysis is used to analyze the shopping habits of consumers. This paper used Market Basket Analysis with the Double Association rule method. Double association rule is used because, a single association rule cannot cover the desired results. The results of this study are recommendations for product layout and product bundling recommendations. These recommendations are implemented to increase customer loyalty so that retail revenue also increases. 71.1% of respondents agreed with the proposed implementation of increasing income and customer satisfaction.

Key words : Customer Relationship Management, Double Association Rule, Product Based, Data Mining, FP-Growth.

1. INTRODUCTION

Retail is all business activities related to selling products or services directly to end consumers [1]. The retail industry is one of the business activities that have very high competition [2]. One of the fast-growing retail industries in Indonesia is retail in a variety of basic needs that are available in Supermarkets. The emergence of many supermarkets resulted in high competition between company management. High competition between companies requires companies to innovate and run specific strategies to compete with other companies. The strategy is carried out to increase revenue where the increase in income comes from the value of transactions made by consumers. The strategy that can be used is a customer relationship management strategy.

The consumer is one of the essential assets of the company. Consumers are essential for the survival of the company [3], [4]. One thing that must be done in a company is to retain existing customers and add new customers with various

services provided by the company. Companies can approach consumers by knowing consumer behavior in the company. One way that can be used to determine consumer habits is by shopping cart or market basket analysis, where it is part of customer relationship management.

Customer Relationship Management (CRM) is a business strategy in which a customer-centric philosophy and culture are applied to marketing, sales, and service activities. Customer-centric is to prioritize consumers from other factors [5]. Customer relationship management is the entire process to build and improve relationships between companies and consumers who give gifts to customers to increase customer satisfaction between companies and consumers increases [6]. One customer relationship management application in retail is to use market basket analysis. Market Basket Analysis (MBA) is a special type of affinity analysis where an interesting event is a product purchase [7].

Transaction data that continues to increase every day will be useless if not used. The transaction data can be used as input data for the Market Basket Analysis process. Market basket analysis is used to determine consumer habits so that by knowing consumer habits can prepare strategies to make consumers more loyal. Many supermarkets have implemented market basket analysis [8], [9]. Market basket analysis produces many business strategies, for example, recommendations for product layout and product bundling combination recommendations to increase consumer impulse buying [10].

In this study, transaction data is used as a data source to overcome sales problems. The data source in this study is transaction data at Niki Laris Supermarkets. The sales problem discussed in this study is only sales optimization within the scope of product layout and promotion strategies in the department's daily needs. The Double FP-Growth algorithm is used to solve problems in product layout and promotional strategy. The Double FP-Growth algorithm is used because the single FP-Growth algorithm cannot produce the desired two outputs and cannot be used in multiple level data with single FP-Growth, so using double FP-Growth.

2. LITERATURE REVIEW

A lot of research on market basket analysis has been done. In 2011 there was a study entitled The Rules of Mine-Oriented Association: An Efficient Approach Based on Profit and

Quantity. This research develops a simple a priori algorithm with the discussion of two new parameters that are contained in the value of u_gain and w_gain . The result of this research is that combining a priori algorithms with u_gain and w_gain can produce association rules that are more effective and can be applied for business development [11].

In 2016, a study entitled Optimization of Rules Selection for Product Bundling Recommendations Through a Combination of Apriori Algorithms and Utility Weight Scores (UWS). In this study, we were studying a priori algorithms and heavy utilities for bundling products [12].

In 2018 there is a study entitled "Market Basket Analysis from Egocentric Videos". This study used a camera mounted on a shopping basket. This study uses the VMBA method to find out which transactions will be used as data sets. In this study successfully studied the habits of consumers in making purchases and ordering purchases [9].

In 2019 a research entitled "Redesign of Layout of Storefronts with a Market Basket Analysis and Activity Relationship Charts" (Case Study of Retail Lawson of Bunda Mulia University) was presented in 2019. This study uses the Basket Market Analysis method to solve the problem of the product display layout at Lawson. This study collects recommendations and research on consumers who know transaction data using the Market Basket Analysis method [13].

In 2019 the research was titled "Negative Mining Regular items from Data Stream". This study uses a new parameter that is the sequence of retrieval and is carried out to find negative mining often in string data [14].

In 2019 there was a study entitled "The Negative Mining Association between Regular Patterns and Hidden Frequencies in the Static Database". This research is looking for association rules on static data [15].

In 2019 there was a study entitled "On Mining Incremental Databases for Regular and Frequent Patterns". The method proposed in this study uses the single window concept that allows for pointer resolution using whichever is selected - transactions contained in the window are used to generate patterns [16].

2.1 Data Mining

Data Mining is the process of discovering interesting patterns and knowledge from large amounts of data. The data sources can include databases, data warehouses, the Web, other information repositories, or data that are streamed into the system dynamically [17].

Data mining can be divided into several stages. The stages are as follows [17]: Data Cleaning, Data Integration, Data Selection, Data Transformation, Mining Data, Pattern Evaluation and Knowledge Presentation.

2.2 Market Basket Analysis

Market Basket Analysis is one of the modeling techniques in data mining to analyze consumer shopping habits by finding associations between several different products, which consumers place in a shopping basket purchased on one

particular transaction [17]. In another sense, market basket analysis is a method of knowing knowledge about the relationships between products, both nominal and categorical products [18].

Market basket refers to the collection of products contained in a single transaction. In this case, the quantity of each item does not affect the process in the market basket analysis. Market basket analysis is only based on different types of products, not on how many of the same products are purchased. The purpose of market basket analysis is to find out the relationship between one item and another item that can be used to assist management decisions in marketing strategies [17].

2.3 Association Rule

Association Rule is a data mining technique for finding patterns of relationships in data or databases of the most popular is Market Basket Analysis (MBA). Association rule is one method that aims to find patterns that often appear among many transactions, where each transaction consists of several products [19]. The idea of the association rule is to examine all possible if-then relationships between products and choose only the most likely (most likely) indicators of the dependency relationship between products. Usually, the term antecedent is used to represent the "if" part and consequent to represent the "then" part in this analysis. Antecedent and consequent are a group of products that are not related together [17]. There are several calculations in the association rules: Support, Confidence and Lift Ratio.

2.4 Support

This stage looks for combinations of products that meet the minimum requirements of the support value in the database. The value of an item's Support is obtained by the formula in (1) and (2) [17]:

$$Support(A) = \frac{\sum Transaction A}{\sum Transaction} \quad (1)$$

$$Support(A, B) = \frac{\sum Transaction (A \cap B)}{\sum Transaction} \quad (2)$$

2.5 Confidence

After all, support is sought, then the association rules that meet the minimum requirements are sought. To find the value of confidence $A \rightarrow B$ obtained by the formula in (3) [17].

$$Confidence = \frac{\sum Transaction (A \cap B)}{\sum Transaction A} \quad (3)$$

2.6 Lift Ratio

A better way to see whether strong association rules are to compare with benchmark values, where we assume the occurrence of products of consequent consequences in a transaction, is independent of the events of the antecedent of an association rule. Or, in other words, if the products in antecedent and consequent are mutually independent. The

Confidence Benchmark formula can be seen in (4)[17]:

$$ConfidenceBenchmark = \frac{\sum Transactionconsequent}{\sum Transaction} \quad (4)$$

Lift Ratio is a comparison between confidence for all rules divided by confidence benchmarks. Lift Ratio values greater than 1 indicate the benefits of these rules. The higher the value of the Lift Ratio, the greater the strength of the association. The formula can be seen in (5)[17]:

$$LiftRatio = \frac{confidence}{benchmarkconfidence} \quad (5)$$

2.7 FP-Growth

FP-Growth algorithm is an algorithm of association rule technique that can be used to determine the set of data that most often appears (frequent product set) in a data set. The development of this a priori algorithm lies in the scanning database and the accuracy of the rules. FP-Growth is more beneficial because it is only done once or twice when scanning the database, while a priori needs to do repeated database scanning. In a priori the accuracy of the rules is higher than FP-Growth, but due to repeated scanning, the speed is slower. However, both algorithms have the same goal of determining frequent product set [20].

Extracting frequent product set using the FP-Growth algorithm will be done by generating a data tree structure or called FP-Tree. The FP-Growth method can be divided into three main stages: The generation phase of the conditional pattern base, The FP-Tree conditional generation stage, The frequent product set search stage.

2.8 Evaluation

Product evaluation is an evaluation of product results, namely identifying and assessing both short- and long-term results. Evaluation of outputs is directed at the program's direct results [21].

3. PROPOSED METHOD

This paper conducts a case study on Niki Laris Supermarkets to determine the layout of new products and determine product bundling. The steps taken in this paper are shown in Figure 1.

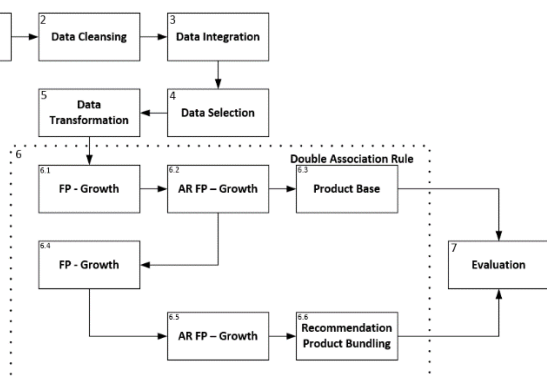


Figure 1: Step Method

Figure 1 is a figure that explains the steps to be taken in this study. The first step is Data Collection, in this step the transaction data collection is carried out from Niki Laris Supermarket with the period August 2018 until August 2019. The next step is Data Cleaning, at this stage the data cleaning is done by editing or removing the wrong value or format. The next step is Data Integration at this stage the process of combining several databases into one. The next step is Data Selection, at this stage the process of selecting data will be used and not used. The next step is Data Transformation, at this stage the format of the data is changed into tabular data so that it can be used as input to the next process. The first step in mining is to use the results of the data transformation process using FP-Growth. The first FP- Growth results will be used as a recommendation for item layout. The association rules formed from the first FP - Growth process will be used as a reference in the second FP- Growth process whose results will be used for product bundling recommendations. Product layout recommendations and product bundling recommendations will be applied and evaluated using the questionnaire method.

4. RESULT AND ANALYSIS

This section will discuss the steps of the study, which were explained in the previous section. The first step is data collection, the data source in this study comes from sales transaction data for the period of August 1, 2018 - August 1, 2019. This data will be processed to produce knowledge that can be used for business strategy development. The data obtained are transaction data including transaction tables, products tables, category tables, subcategories tables and transaction details tables. The data obtained amounted to 72,645 transactions and 11,164 products.

4.1 Data Cleaning

This stage is not carried out because the data taken is transactional data that has already been completed.

4.2 Data Integration

This stage is a process of merging data from several supporting sources, but because the data source is used only from one database, then this stage is not carried out.

4.3 Data Selection

This stage the data selection process is carried out, this process is carried out to sort out the data used and not used. After this selection process, the amount of data used amounted to 40,940 transactions and 3,339 items. 3,339 items are categorized into 6 categories, these 6 categories can be seen in Table 1.

Table 1: Category Product

No	Category	Total
1	Herbs	205
2	Processed Food	127
3	Snack	1455
4	Processed Drink	595
5	Instant Drink	483
6	Medicines and Vitamins	474

4.4 Data Transformation

This stage is carried out the process of changing transaction data into tabular data, is shown in Table 2 and Table 3.

Table 2: Transaction Data

No	Id Transaction	Id Products	Qty	Price
1	317020880012	7693	1	13500
2	11703160033	7693	2	27000
3	11703160033	7693	9	40500

Table 3: Tabular Data

No	Id Transaction	Herbs	Processed Food	Snack	Processed Drink	Instant Drink	Medicines and Vitamins
1	317020880012	1	1	0	0	0	0
2	11703160033	1	1	0	0	0	1
3	11703160033	1	0	0	0	0	1

4.5 First FP – Growth

This stage an analysis using FP-Growth is used to see the interrelationships between product categories. FP-Growth requires a reference value that must exist, namely confidence and support. This stage using threshold support 22% and confidence 54%, this value is obtained by trial and error. The results of this stage produce 5 association rules which are shown in Table 4.

Table 4: Association Rule FP-Growth

No	Antecedent	Consequent	Support	Confidence	Lift Ratio
1	Tea	Instant Drinks	0.12	0.56	1.05
2	Instant Drinks	Sugar	0.31	0.57	1.3
3	Sugar	Instant Drinks	0.31	0.75	1.3

4.6 AR First FP-Growth

Five association rules that are formed have lift ratio values that are close to one or more than one, it means that the association rules that are formed are positive. The association rules will be used for recommendations on the layout of the products.

4.7 Products Layout Recommendation

Products Layout Recommendation is obtained from the association rules that have been formed in the previous process. Item layout recommendations are shown in Figure 2.

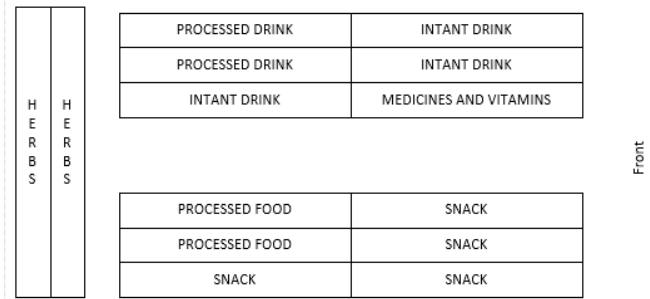


Figure 2: Products Layout Recommendation

4.8 Second FP-Growth

This stage, the second FP-Growth is carried out by inputting the results of the first FP-Growth. Of the five association rules formed, there are only three combinations and 4 product categories, namely: Instant Drink and Processed Drink, Snack and Instant Drink, Processed Food and Snack, and 4 product categories, namely Instant drink, snack, processed drink, and processed food. This stage will be done once for each association rule that is formed to see the new association rules from the subcategory of goods.

The results of the rules for the association of Instant drink and processed drinks with threshold support of 12% and confidence of 56% are shown in Table 5.

Table 5: Association Rule FP-Growth Subcategory

No	Antecedent	Consequent	Support	Confidence	Lift Ratio
1	Instant Drink	Processed Drink	0.28	0.54	1.4
2	Snack	Instant Drink	0.35	0.57	1.1
3	Processed Food	Snack	0.22	0.58	0.9
4	Instant Drink	Snack	0.35	0.68	1.2
5	Processed Drink	Instant Drink	0.28	0.73	1.7

The results of the snack and instant drink association rules with a 15% threshold support and 63% confidence are shown in Table 6.

Table 6: Association Rule FP-Growth Subcategory

No	Antecedent	Consequent	Support	Confidence	Lift Ratio
1	Candy	Snack	0.15	0.63	0.746
2	Mineral Water	Snack	0.29	0.8	0.949
3	Instant Milk	Snack	0.31	0.84	0.993
4	Instant Drink	Snack	0.23	0.88	1.041

The results of the association rules for processed food and snack drinks with a 30% threshold support and 74% confidence are shown in Table 7.

Table 7: Association Rule FP-Growth Subcategory

No	Antecedent	Consequent	Support	Confidence	Lift Ratio
1	Sardines	Noodle	0.3	0.92	1.1
2	Sardines	Snack, Noodle	0.3	0.92	1.2
3	Snack, Sardines	Noodle	0.3	0.92	1.1
4	Noodle	Snack	0.59	0.97	0.9
5	Sardines	Snack	0.4	1	1
6	Noodle, Sardines	Snack	0.3	1	1

4.9 AR Second FP-Growth

Three association rules are formed from the rules of the association of Instant drink and processed drinks are shown in Table 5, four association rules that are formed from the rules of the association of snacks are shown in Table 6 and instant drinks and six association rules that are formed from the rules of the association of processed food and snacks are shown in Table 7.

4.10 Recommendation Products Bundling

This stage, product bundling recommendations will be based on the association rules that have been generated in the second phase of FP-Growth. This product bundling recommendation can be applied to the cross-selling method. Product bundling recommendations are shown in Table 8.

Table 8: Recommendation Product Bundling

No	Recommendation
1	Tea, Sugar, Instant Drinks
2	Tea, Sugar
3	Instant Drink, Snack
4	Noodle, Snack
5	Noodle, Snack, Sauce
6	Noodle, Sardines

4.11 Evaluation

Both recommendations will be implemented in Niki Laris Supermarkets and will be evaluated with a questionnaire for consumers and internal employees. The questions in the questionnaire are shown in Table 9.

Table 9: Questionnaire Evaluation

No	Questionnaire	Filler
1	Does implementing the products layout recommendation reduce consumers in asking the location of the product?	Internal (10)
2	Is there an increase in sales from promotional recommendations in accordance with product bundling recommendations?	Internal (10)
3	Is there an increase in daily income after applying the two recommendations from research?	Internal (10)
4	Does this recommendation have a big impact in Niki Laris Supermarkets?	Internal (10)
5	Does this new products layout help you to shop?	Consumers (50)
6	Does the bundling product promotions / shopping package promotions fit your needs?	Consumers (50)
7	Are you interested in bundling product promotions / shopping package promotions?	Consumers (50)

The questionnaire was distributed to 10 internal employees and 50 consumers with the results shown in Figure 3.

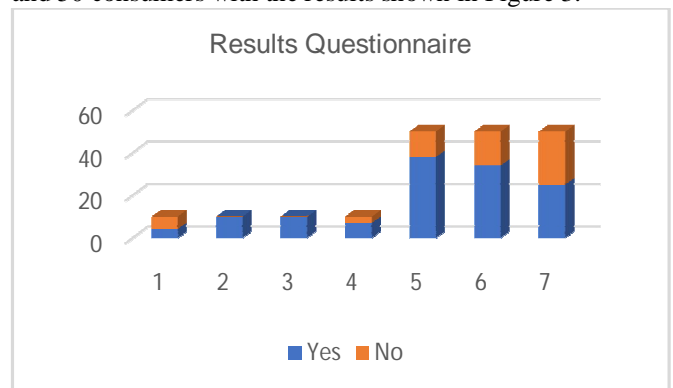


Figure 3: Questionnaire

Figure 3 is a picture showing the answers to the questionnaire. Questionnaires number 1-4 were filled internally with the results of 77.5% answering yes and 22.5% answering no. Questionnaires number 5-7 were filled by consumers with the results, 64.7% answered yes, and 35.3% answered no. From these two results, it can be concluded that 71.1% agreed to the implementation of the results of this study, which had a positive impact.

5. CONCLUSION

Based on the results of the rules established association can be used as recommendations for the layout of goods and product bundling recommendations as a step to implement customer relationship management in the retail department. As many as 71.1% of respondents agreed with the proposed implementation.

Based on the analysis that has been done, there are some deficiencies that can be developed in other studies. Therefore, for further research it is recommended to use more complex transaction data, the use of more data can improve the accuracy of association rules, broader range of goods categories and add new parameters to provide better results.

ACKNOWLEDGEMENT

We thank Mr. Sujito as the management of Niki Laris Supermarkets for giving us the freedom to conduct research in this supermarket. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

1. P. Kotler and G. Amstrong. *Manajemen Pemasaran Jilid 2*, Jakarta: Penerbit Erlangga, 2013.
2. I. Surjandari and A. C. Seruni. **Design of Product Placement Layout in Retail Shop Using Market Basket Analysis**, *MAKARA Technol. Ser.*, vol. 9, no. 2, pp. 43–47, 2010. <https://doi.org/10.7454/mst.v9i2.379>
3. C. Irawan. *Secangkir Kopi Untuk Sahabat Customer*, Jakarta: PT Gramedia Pustaka Utama, 2013.
4. F. Rangkuti. *Customer Care Excellence: Meningkatkan Kinerja Perusahaan Melalui Pelayanan Prima Plus Analisis Kasus Jasa Raharja*, Jakarta: Gramedia Pustaka Utama, 2017.
5. E. K. Budiardjo, F. Irwiensyah, and D. Obstetri. **Analisis Fitur Crm Untuk Meningkatkan Kepuasan Pasien - Berbasis Pada Framework of Dynamic Crm Studi Kasus : Departemen Obstetri Dan Ginekologi Fkui - Rscm Jakarta**, *Perspective*, 2008.
6. Z. Soltani and N. J. Navimipour. **Customer Relationship Management Mechanisms: A Systematic Review of The State Of The Art Literature And Recommendations For Future Research**, *Comput. Human Behav.*, vol. 61, pp. 667–688, 2016. <https://doi.org/10.1016/j.chb.2016.03.008>
7. M. Dhanabhakayam and M. Punithavalli. **A Survey on Data Mining Algorithm For Market Basket Analysis**, *Glob. J. Comput. ...*, vol. 11, no. 11, pp. 22–28, 2011.
8. S. C. Bilqisth and S. Astuti. **Analisis Pola Pembelian Konsumen Dengan Algoritma Apriori pada IndomaretIndraprasta Semarang**, *Tek. Inform.*, 2015
9. V. Santarcangelo, G. M. Farinella, A. Furnari, and S. Battiato. **Market Basket Analysis From Egocentric Videos**, *Pattern Recognit. Lett.*, vol. 112, pp. 83–90, 2018.

10. J. Bermúdez, K. A. Rodríguez, and A. G. Abad. **Layout Optimization and Promotional Strategies Design in a Retail Store based on a Market Basket Analysis**, *Int. Multi -Conference Eng. Educ. Technol.*, pp. 20–22, 2016. <https://doi.org/10.18687/LACCEI2016.1.1.307>
11. P. S. Sandhu, D. S. Dhaliwal, and S. N. Panda. **Mining Utility-Oriented Association Rules: An Efficient Approach Based On Profit And Quantity**, *Int. J. Phys. Sci.*, vol. 6, no. 2, pp. 301–307, 2011.
12. I. Fahmi, H. Suyono, and M. Sarosa. **Optimasi Seleksi Aturan Untuk Rekomendasi Bundling Produk Melalui Kombinasi Algoritma Apriori dan Utility Weighted Score (UWS)**, *EECCIS J.*, vol. 10, no. 2, pp. 59–62, 2016.
13. F. R. Wilujeng, W. Wu, and F. Nurprihatin. **Perancangan Ulang Tata Letak Etalase Barang Dengan Metode Market Basket Analysis Dan Activity Relationship Chart (Studi Kasus Retail Lawson Universitas Bunda Mulia)**, *Proceeding SENDI_U*, 2019.
14. N. P. Kumar et al. **Mining Negative Frequent Regular Itemsets From Data Streams**, *Int. J. Emerg. Trends Eng. Res.*, vol. 7, no. 8, pp. 85–98, 2019. <https://doi.org/10.30534/ijeter/2019/02782019>
15. P. K. NVS. **Mining Negative Associations between Regular and Frequent Patterns hidden in Static Databases**, *Int. J. Emerg. Trends Eng. Res.*, 2019.
16. N. V. S. P. Kumar, J. K. R. Sastry, and K. R. Sekhara Rao. **On Mining Incremental Databases for Regular And Frequent Patterns**, *Int. J. Emerg. Trends Eng. Res.*, vol. 7, no. 9, pp. 291–305, 2019. <https://doi.org/10.30534/ijeter/2019/12792019>
17. J. Han, M. Kamber, and J. Pei. *Data Mining: Concepts and Techniques*, San Francisco: Morgan Kaufmann Publishers, 2012.
18. Raorane AA, Kulkarni RV, and Jitkar BD. **Association Rule-Extracting Knowledge Using Market Basket Analysis**, *Res. J. Recent Sci.*, vol. 1, no. 2, pp. 19–27, 2012.
19. M. Kaur and S. Kang. **Market Basket Analysis: Identify the Changing Trends of Market Data Using Association Rule Mining**, *Procedia Comput. Sci.*, vol. 85, no. Cms, pp. 78–85, 2016. <https://doi.org/10.1016/j.procs.2016.05.180>
20. M. J. Zaki and W. Meira, Jr. *Data Mining and Analysis, in Data Mining and Analysis*, Cambridge University Press, 2018.
21. R. Ananda and M. hu. Tien Rafida. **Pengantar Evaluasi Program pendidikan**. 2017.