The Use of Apriori Algorithm in the Formation of Association Rule at Lotteria Cibubur

Ovi Liansyah  
STMIK Nusa Mandiri Jakarta  
Jakarta, Jl. Melati 3, RT.002/RW.012, Jatiwaringin, Kec. Pondokgede, Kota Bks, Jawa Barat  
ovi_liansyah11@gmail.com

Henny Destiana  
UBSI Jakarta  
Jakarta, Jl. Kamal Raya No.18, RT.6/RW.3, Cengkareng, Kota Jakarta Barat  
Henny_hnd@bsi.ac.id

Submitted: Mar 11, 2020  
Accepted: Mar 21, 2020  
Published: Apr 1, 2020

Abstract — Lotteria as one of the franchises that produce sales data every day, has not been able to maximize the utilization of that data. The sale data storage is still not optimal. By utilizing sales transaction data that have been stored in the database, the management can find out the menus purchased simultaneously, using the association rule. Namely, data mining techniques to find the association rules of a combination of items. The process of searching for associations uses the help of apriori algorithms to produce patterns of the combination of items and rules as important knowledge and information from sales transaction data. By using the minimum support parameters, the minimum and the month period of the sales transaction to find the association rules, the data mining application generates association rules between items in April 2019, where consumers who buy hot / ice coffee will then buy together with support of 16% and 100% confidence. Knowing which menu products or items are the most sold, thus lotteria Cibubur can develop a sales strategy to sell other types of menu products by examining the advantages of the most sold menu with other menus and can increase the stock of menu ingredients.

Keywords — Data mining, Menu, Apriori Algorithm, Database, association rule

I. INTRODUCTION

An application is needed to analyze the market basket of drug sales transaction data using data mining as a data analysis technique that can help pharmacies obtain knowledge in the form of sales patterns within a certain month period. Data mining applications are built using linear sequential processes with the PHP programming language and MySQL database. The algorithm used as the main process of market basket analysis is apriori algorithm using the parameters of minimum support, minimum confidence, and the period of months of sales transactions to find the association rules. (Irfrani, 2019)

In increasing the company's turnover can be done using the Data Mining process, one of which is to use apriori algorithm. With the apriori algorithm, association rules can be found which can later be used as a pattern of purchasing goods by consumers, this study uses a data repository of 209 records consisting of 23 transactions and 164 attributes. From the results of this study, the item named CREAM CUPID HEART COAT HANGER is the product most often bought by consumers. (Putra et al., 2019)

Sales of OPPO brand mobile electronic products have not used data mining implementation, where all goods that are already in stock in the showroom must be sold, all cannot be purchased at the center. To avoid the buildup of stock of mobile phones that are less interested and know what type of opposition brands with the most sales in the opposition SDC store requires an apriori algorithm. This can be known using apriori algorithm which is part of data mining. (Kanti & Indrajit, 2017)

By using a priori algorithm, it can produce a combination pattern of 17 (seventeen) rules with a support value of 70% and the highest confidence value of the 17 (seventeen) rules of 93% contained in the Lock Broken → Disscolour rule. (Pahlevi & Sugandi, 2019)
Evaluation and evaluation of lecturers' performance conducted at STT Harapan is still not optimal so that the level of awareness of each lecturer to improve their performance is also not optimal. This study aims to apply data mining algorithm C4.5 in determining the track record of lecturer performance based on 3 input attributes (publication, service and teaching) and 1 output attribute that has "less" "sufficient" and "high" performance. The results of the study are decision trees and rules that provide information on the results of the evaluation of the performance of STT Harapan Medan lecturers in implementing the tri dharma of tertiary institutions. (Rismayanti et al., 2019).

II. LITERATURE REVIEW

Data Mining discuss extracting or gathering useful information from data sets. Information that is usually collected is hidden patterns in data, relationships between data elements, or modeling for the purpose of forecasting data. (Santoso et al., 2016)

Association Rule Mining is a data mining technique for finding associative rules between combinations of items. An example of an associative rule from a purchase analysis at a supermarket is identifying how likely a customer is to buy bread together with milk. (Saw, 2019)

Apriori algorithm is the most famous algorithm for finding high frequency patterns. Apriori algorithm is divided into several stages called narration or pass. Formation of itemset candidates, candidates are formed from a combination (k1) -itemset obtained from the previous iteration. One way of apriori algorithm is to prune candidates whose subsets containing k-1 items are not included in a high-frequency pattern with a length of k-1. (Saw, 2019)

This stage is a research sequence design from beginning to end in conducting research:

1. First study
   The initial study of this research is by looking from studying the problem to be investigated. Then determine the scope of the problem, the background of the problem, and study some literature related to the problem and how to find solutions to the problem

2. Data collection
   The data collection of this study the authors conducted interviews, observations and documentation. To find out the information needed, the writer collected sales data in the April 2019 period in Lotteria.

3. Data Processing with Data mining
   At this stage the data processing first identify the problems often faced by the restaurant then describe these problems in order to obtain a solution. The next stage is analyzing the problem using data mining with apriori algorithm to get the results as a goal to be achieved then it can be used by the restaurant as knowledge in increasing product sales.

4. Analysis of Results
   At this stage the authors analyze the results using the Tanagra 1.4 application and apriori algorithm in connecting data to be tested. (Yanto & Khoiriah, 2015)

Data Mining has been implemented in various fields, including business or trade, education and telecommunications. In business, for example, the results of implementing data mining using the Apriori algorithm can help business people in decision making policies to what is related to inventory. For example the importance of the inventory system in a Pharmacy and what types of goods are the top priority that must be in stock to anticipate vacancy of goods. Because of the lack of stock of goods can affect customer service and income Pharmacy. Therefore the availability of various types of medical devices at the Pharmacy as one of the suppliers of medical equipment, absolutely to support the smooth distribution to consumers, so that the activity customer service going well. (King et al., 1966)

How to find books that are bought together, can be used association rule, which is a data mining technique to find the association rules of a combination of items. The association search process uses the help of a priori algorithms to produce patterns of combination of items and rules as important knowledge and information from sales transaction data. The results of this study are in the form of applications to analyze spending patterns where the resulting pattern can be used as recommendations in determining sales strategies by Gramedia. (Listriani et al., 2018)
III. PROPOSED METHOD

A. Formation of 1 Itemset
With a minimum support value of 15%, the support value of 1 item is obtained using the following formula:

\[ \text{Support}(A, B) = \frac{\sum \text{Transaction Contains } A}{\sum \text{Transaction}} \times 100\% \]

The following is the calculation of the formation of 1 itemset:

\[ S(\text{Float}) = \frac{\sum \text{Float}}{30} = \frac{22}{30} \times 100\% = 73\% \]

\[ S(\text{Ice Lemon Tea}) = \frac{\sum \text{Ice Lemon Tea}}{30} \]
\[ = \frac{13}{30} \times 100\% = 43\% \]

\[ S(\text{Buldak Burger}) = \frac{\sum \text{Buldak Burger}}{30} \]
\[ = \frac{1}{30} \times 100\% = 3\% \]

\[ S(\text{Cappucino}) = \frac{\sum \text{Cappucino}}{30} \]
\[ = \frac{13}{30} \times 100\% = 43\% \]

\[ S(\text{Hot / Ice Coffe Latte}) = \frac{\sum \text{Hot / Ice Coffe Latte}}{30} \]
\[ = \frac{3}{30} \times 100\% = 10\% \]

\[ S(\text{Ice Coco}) = \frac{\sum \text{Ice Coco}}{30} \]
\[ = \frac{2}{30} \times 100\% = 6\% \]

\[ S(\text{Chicken Puas A}) = \frac{\sum \text{Chicken Puas A}}{30} \]
\[ = \frac{17}{30} \times 100\% = 56\% \]
The support value of the 1 item that has been described can be seen in the table below:

### TABLE 2. Value of Support 1 Itemset

| No | Menu Table           | Itemset | Amount | Support |
|----|----------------------|---------|--------|---------|
| 1. | Float                | 22/30   |        | 73%     |
| 2. | Ice Lemon Tea       | 13/30   |        | 43%     |
| 3. | Cappucinno          | 13/30   |        | 43%     |
| 4. | Chicken Puas A      | 17/30   |        | 56%     |
| 5. | Hot/Ice Coffe Latte | 5/30    |        | 16%     |
| 6. | Ico Coco            | 5/30    |        | 16%     |
| 7. | Chicken Puas A      | 17/30   |        | 56%     |

With a minimum support of 15%, then the combination of 1 itemset that does not meet the minimum support will be removed in the following table:

### TABLE 3.Minimum Support 1 Itemset

| No | Menu Table           | Itemset | Amount | Support |
|----|----------------------|---------|--------|---------|
| 1. | Float                | 22/30   |        | 73%     |
| 2. | Ice Lemon Tea       | 13/30   |        | 43%     |
| 3. | Cappucinno          | 13/30   |        | 43%     |
| 4. | Chicken Puas A      | 17/30   |        | 56%     |
| 5. | Hot/Ice Coffe Latte | 5/30    |        | 16%     |
| 6. | Ico Coco            | 5/30    |        | 16%     |

### B. Formation of 2 Itemset

Formation of Combination

The support value of 2 items is obtained by the following formula:

\[
\text{Support}(A, B) = \frac{\sum \text{Transaction Contains } A \text{ dan } B}{\sum \text{Transaction}} \times 100\%
\]

The following is a calculation of 2 item sets:
The support value of the 2 items that have been obtained from the description above can be seen in the table below:
### TABLE 4. Value of Support 2 Itemset

| No | Menu Table | Amount | Support |
|----|------------|--------|---------|
| 1. | Float, Ice Lemon Tea | 8/30 | 26% |
| 2. | Float, Cappucino | 9/30 | 30% |
| 3. | Float, Chicken Puas A | 13/30 | 43% |
| 4. | Float, Hot/Ice Coffe | 5/30 | 16% |
| 5. | Float, Orange Juice | 1/30 | 3% |
| 6. | Ice Lemon Tea, Cappucino | 3/30 | 10% |
| 7. | Ice Lemon Tea, Chicken Puas A | 5/30 | 16% |
| 8. | Ice Lemon Tea, Hot/Ice Coffe | 2/30 | 6% |
| 9. | Ice Lemon Tea, Orange Juice | 2/30 | 6% |
| 10. | Cappucino, Chicken Puas A | 6/30 | 20% |
| 11. | Cappucino, Hot/Ice Coffe | 1/30 | 3% |
| 12. | Cappucino, Orange Juice | 2/30 | 6% |
| 13. | Chicken Puas A, Hot/Ice Coffe | 2/30 | 6% |
| 14. | Chicken Puas A, Orange Juice | 1/30 | 3% |
| 15. | Hot/Ice Coffe, Orange Juice | 0/30 | 0% |

With a minimum support of 15%, the combination of 2 itemset that does not meet the minimum support will be removed in the following table:

### TABLE 5. Minimum Support 2 Itemset

| No | Menu Table | Amount | Support |
|----|------------|--------|---------|
| 1. | Float, Ice Lemon Tea | 8/30 | 26% |
| 2. | Float, Cappucino | 9/30 | 30% |
| 3. | Float, Chicken Puas A | 13/30 | 43% |
| 4. | Float, Hot/Ice Coffe | 5/30 | 16% |
| 5. | Cappucino, Chicken Puas A | 6/30 | 20% |

C. Formation of 3 Itemset

Formation Combination The support value of 3 items is obtained by the following formula:

$$\text{Support} (A, B, C) = \frac{\sum_{\text{Transaction}} \text{Cont}_A, B \text{dan} C}{\sum_{\text{Transaction}}} \times 100\%$$

The following is a calculation of 3 item sets:

$$S(\text{Float, Ice Lemon Tea, Cappucino}) = \frac{\sum \text{Float, Ice Lemon Tea, Cappucino}}{\sum 30} \times 100\% = 6\%$$

$$S(\text{Float, Ice Lemon Tea, Chicken Puas A}) = \frac{\sum \text{Float, Ice Lemon Tea, Chicken Puas A}}{\sum 30} \times 100\% = 10\%$$

$$S(\text{Float, Ice Lemon Tea, Hot / Ice Coffe}) = \frac{\sum \text{Float, Ice Lemon Tea, Hot / Ice Coffe}}{\sum 30} \times 100\% = 6\%$$

The support value of the 3 items obtained can be seen in the table below:

### TABLE 6. Itemset Support Value

| No | Menu Table | Amount | Support |
|----|------------|--------|---------|
| 1. | Float, Ice Lemon Tea, Cappucino | 2/30 | 6% |
| 2. | Float, Ice Lemon Tea, Chicken Puas A | 3/30 | 10% |
| 3. | Float, Ice Lemon Tea, Hot/Ice Coffe | 2/30 | 6% |
| 4. | Cappucino, Chicken Puas A, Hot/Ice Coffe | 0/30 | 0% |
Because of the formation of 3 itemset none meet the 15% support value, then the combination of 2 itemset that meets the formation of the Association. After all the high frequency patterns have been found, then the association rules are found that meet the minimum requirements for confidence by calculating the confidence of associative rules A → B. Minimum Confidence = 80%. Confidence value from the rules A → B is obtained by the following formula:

\[ \text{Confidence}(A, B) = \frac{\sum \text{Numb transaction contains A \& B}}{\sum \text{TOTAL Transaction A}} \times 100\% \]

The following is a calculation of 2 item sets:

\[ S(\text{Float, Ice Lemon Tea}) = \frac{\sum \text{Float, Ice Lemon Tea}}{\sum 22} = \frac{8}{22} \times 100\% = 36\% \]

\[ S(\text{Ice Lemon Tea, Float}) = \frac{\sum \text{Ice Lemon Tea, Float}}{\sum 13} = \frac{9}{13} \times 100\% = 69\% \]

\[ S(\text{Float, Cappucino}) = \frac{\sum \text{Float, Cappucino}}{\sum 22} = \frac{8}{22} \times 100\% = 36\% \]

\[ S(\text{Cappucino, Float}) = \frac{\sum \text{Cappucino, Float}}{\sum 13} = \frac{8}{13} \times 100\% = 61\% \]

\[ S(\text{Float, Chicken Puas A}) = \frac{\sum \text{Float, Chicken Puas A}}{\sum 22} = \frac{13}{22} \times 100\% = 59\% \]

\[ S(\text{Chicken Puas A, Float}) = \frac{\sum \text{Chicken Puas A, Float}}{\sum 17} = \frac{13}{17} \times 100\% = 76\% \]

Confidence value of 2 items obtained can be seen in the table below:

| No | Menu Table | Amount | Confidence |
|----|------------|--------|------------|
| 1  | Float, Ice Lemon Tea, Cappucino | 2/30   | 6%         |
| 2  | Float, Ice Lemon Tea, Chicken Puas A | 3/30 | 10%        |
| 3  | If you buy Float, you will buy Ice Lemon Tea | 8/22 | 36%        |
| 4  | If you buy Ice Lemon Tea, you will buy Float | 9/13 | 69%        |
| 5  | If You Buy Float, You Will Buy Cappuccino | 8/22 | 36%        |
| 6  | If you buy cappuccino Then it will buy Float | 8/13 | 61%        |
| 7  | If You Buy A Float Then You Will Buy A Satisfied Chicken | 13/22 | 59%        |
| No | Itemset                          | Amount | Confidence |
|----|---------------------------------|--------|------------|
| 8. | If you buy cappuccino Then it will buy Float | 13/17  | 76%        |
| 9. | If You Buy A Float Then You Will Buy A Satisfied Chicken | 5/22   | 22%        |
| 10.| If You Buy Hot / Ice Coffe Then You Will Buy Float | 5/5    | 100%       |
| 11.| If You Buy Cappuccino Then You Will Buy A Satisfied Chicken A | 6/13   | 46%        |
| 12.| If You Buy A Satisfied Chicken A, You Will Buy Cappucino | 6/17   | 35%        |

V. CONCLUSION AND SUGGESTION

A. Conclusions

Based on the discussion that has been done with the Apriori Algorithm and the testing done with the Tanagra application, the writer draws several important conclusions. The conclusions are as follows:

1. Data Mining by using apriori algorithm can be implemented to analyze the needs of the company, by using a product sales transaction database because it can find the trend of itemset combination patterns so that it can be used as very valuable information in making decisions to prepare stock of any type of product needed later.

2. Application of apriori algorithm in data techniques Mining is very efficient and can accelerate the process of forming trends in the combination of itemset patterns of product sales at Lotteria Store Cibubur, namely with the highest support and confidence is the combination of Hot products / Ice and Float are 16% and 100%. This means that Hot / Ice Coffe and Float products are the most popular products by consumers.

B. Suggestion

As for the suggestions submitted so that future research can be better. The suggestions are as follows:

1. For future research, the object of research can be expanded again.
2. The research sample used in future research must be more.
3. Distributing questionnaires to respondents to obtain the data needed in research can be reproduced.
4. Most product sales sold at Cibubur Lotteria Store can be determined using apriori algorithm, by looking at products that meet the minimum support and minimum confidence which products are the most sold, there are difficulties if the data is processed in large quantities and will require time long enough.
5. The application of the Apriori Algorithm is very practical but it needs to be compared with other algorithms, to test the extent to which the Apriori Algorithm is still reliable for processing and finding patterns of relationships (associations) between items in large-scale databases.
6. In using the Apriori Algorithm, an understanding of Association and Data Mining rules and how to run Tanagra software applications is needed.
7. The application of the Apriori Algorithm is very effective but it needs to be compared with other algorithms, to test the extent of the Apriori Algorithm it is still reliable to look for association rules in the data mining process itself.

REFERENCES

Irfiani, E. (2019). Application of Apriori Algorithms to Determine Associations in Outdoor Sports Equipment Stores. *SinkrOn*, 3(2), 218. https://doi.org/10.33395/sinkron.v3i2.10089

Kanti, S., & Indrajit, R. E. (2017). Implementasi Data Mining Penjualan Handphone Oppo Store Sdc Tangerang Dengan Algoritma Appriori. *Seminar Nasional Sains Dan Teknologi, November*, 1–2.

King, D. G., Young, W. E. V., Clarke, A. J., Cain, A. J., & Dimbleby, G. W. (1966).
The Lanhill Long Barrow, Wiltshire, England: An Essay in Reconstruction. *Proceedings of the Prehistoric Society*, 32, 73–85. https://doi.org/10.1017/S0079497X00014341

Listriani, D., Setyaningrum, A. H., & Eka, F. (2018). PENERAPAN METODE ASOSIASI MENGGUNAKAN ALGORITMA APRIORI PADA APLIKASI ANALISA POLA BELANJA KONSUMEN. (Studi Kasus Toko Buku Gramedia Bintaro). *Jurnal Teknik Informatika*, 9(2), 120–127. https://doi.org/10.15408/jti.v9i2.5602

Pahlevi, O., & Sugandi, A. (2019). Penerapan Algoritma Apriori Dalam Pengendalian Kualitas Produk. *Jurnal & Penelitian Teknik Informatika*, 3.

Putra, J. L., Raharjo, M., Sandi, T. A. A., Ridwan, R., & Prasetyo, R. (2019). Implementasi Algoritma Apriori Terhadap Data Penjualan Pada Perusahaan Retail.

Saw, B. M. (2019). Penerapan algoritma apriori dalam pembentukan association rule di lotteria cibubur.