

PAPER • OPEN ACCESS

Study of application of data mining market basket analysis for knowing sales pattern (association of items) at the O! Fish restaurant using apriori algorithm

To cite this article: Yusuf Kurnia *et al* 2019 *J. Phys.: Conf. Ser.* **1175** 012047

View the [article online](#) for updates and enhancements.

You may also like

- [-ALLYL COMPLEXES OF THE TRANSITION METALS](#)
M I Lobach, B D Babitskii and V A Kormer
- [-Isotopic Exchange of Hydrogen in Aromatic Complexes of the Transition Metals](#)
V N Setkina and D N Kursanov
- [-Mass Spectrometry of Transition Metal - Complexes](#)
P E Gaivoronskii and N V Larin



The Electrochemical Society
Advancing solid state & electrochemical science & technology

241st ECS Meeting

Vancouver, BC, Canada. May 29 – June 2, 2022

ECS Plenary Lecture featuring
Prof. Jeff Dahn,
Dalhousie University

Register now!

The banner features the ECS logo, a 'Register now!' button with a checkmark, a photo of Prof. Jeff Dahn, and a background image of the Science World building in Vancouver.

Study of application of data mining market basket analysis for knowing sales pattern (association of items) at the O! Fish restaurant using apriori algorithm

Yusuf Kurnia^{1*}, Yohanes Isharianto², Yo Ceng Giap³, Aditiya Hermawan⁴, Riki⁵

Jl. Imam Bonjol No. 41 Karawaci Ilir-Tangerang 15115

*yusuf.kurnia@ubd.ac.id

Abstract. The development of the food and beverage culinary industry is growing very rapidly. Making food and beverage business owners, especially restaurants, have to make the right decision to stay in a very strong competition, restaurant owners must be ready to always innovate and remain to be able to meet consumer needs through products that can attract customers and determine strategies promotions that can boost sales. Stored transaction data has information that can be extracted by data mining techniques, for example knowing the pattern of sales in purchases by consumers. Information about sales patterns can be used by O! Fish restaurants to create more potential promotional strategies to boost sales by referring to items (menus) that are often purchased together. . To be able to find out the purchase patterns by consumers simultaneously, knowing what products are often purchased simultaneously can be used data mining techniques using a priori algorithms. A priori algorithm is used to generate association rules. Information about the association's rules in purchasing items (menus) by consumers can be used by O! Fish restaurants to create more potential promotional strategies to boost sales by referring to a combination of items that are often purchased simultaneously. Later the results of this study are in the form of a website-based application to analyze purchasing patterns (item association rules) by consumers where the purchase pattern can be used as recommendations in determining the promotion development strategy for O! Fish restaurants.

1. Introduction

The development of the business world in the restaurant sector is happening very fast. This is marked by the emergence of many restaurants with various characteristics or different concepts. "The growth of food and beverage businesses is still recorded as high growth in various parts of the world. With the rapid growth of restaurant businesses, there is now competition among restaurant business owners, this is what makes restaurant owners have to prepare well and mature competition from restaurant concepts, product innovation, customer convenience, market share to the operational system properly. Therefore, to be able to run restaurant operations, restaurant owners must consider developing an operational system in depth to be able to survive and benefit according to the restaurant's objectives.

One of the causes of the development of business in the restaurant sector is that at present the business in the field of food and beverages, especially this restaurant, promises huge profits and sees that human needs will eat and drink as basic needs that must be fulfilled to survive. This is an opportunity that can be used as a profitable business opportunity. No wonder there are so many



restaurants with various concepts, different operational systems, creations with different product innovations that represent what restaurants have to offer to attract customers or concepts that can be shown to consumers.

Seeing the importance of preparation or building a good system to be able to run operations while meeting operational objectives in the business world competition requires a strategy and system that can support and improve restaurant goals. Restaurant owners must still pay attention to various factors in formulating this strategy, for example by paying attention to the products to be sold, product innovations, knowing the buying interest of consumers in marketing and product promotion. It is not easy to be able to grow buying interest before consumers finally decide to buy a product. To be able to overcome this is by promoting a product with promotion causing people who were previously not interested in buying a product will be interested in buying a product and trying to make a purchase. Determining a good promotion system will produce added value for restaurant owners [1]

O! Fish restaurants in practice in carrying out restaurant strategies and operations are often found in menus that are rarely ordered by consumers and in the end the menu cannot have a positive impact and does not generate profits for the restaurant. This happens because the restaurant still uses a manual promotion system and it is less effective in making promotions to boost sales, as a result of this will be a disadvantage for the restaurant because the ingredients of the product over time will decrease in terms of product enjoyment. This will also affect the lack of preparation in the supply of raw goods from the product and affect the recording of data from inventory. If there is a lack of preparation between material stock and sales information, the operational process will be hampered. This will certainly hamper the process of selling restaurants. In the case of O! Fish restaurant promotion, in practice, only do promotion through a simple method using the "Upselling" system, this system is a system that sets out for menus that are not ordered by customers who are then given menu points that are targeted by the parties server / waiters can be promoted to customers so that they can increase the value of the menu that is not sold and become a bonus for the server / waiters who can meet the specified points. Using this method is less effective in promoting customers. And in its implementation this is less run and requires a long time.

To make the right decision in determining the sales strategy and developing the right target promotion is knowing the tastes of the consumer, and knowing the consumer's buying habits regarding what products (food and drinks) are often purchased, knowing the product purchase patterns by consumers, which can be observed through purchase transaction data. Stored transaction data has information that can be extracted to produce new data or data patterns by using data mining. For example information to find out the sales pattern regarding the menu association rules. A priori algorithm or often called market basket analysis (market basket analysis) is a method that can be used to analyze the behavior of consumers' buying patterns to produce association rules with the "if then" pattern. Aims to find out products (food and drinks) that are often purchased simultaneously from transaction data. Sometimes from data processing methods that simply cannot provide effective results because of the large volume of data processed and the difficulty of seeing associations between product sales in this case food and drinks with one another. Thus there is a need for a system that can provide fast and precise information. The use of information and knowledge contained in the amount of data is called "data mining"

2. Related Works

Data mining is a process that uses statistical techniques, mathematics, artificial intelligence, and machine learning to extract and identify useful information and related knowledge from various large databases [1].

Association Rule Mining is a data mining technique to find associative rules between a combination of items. Because association analysis is well-known because its application is for analyzing shopping cart contents, association analysis is also often called market basket analysis to find association rules that meet the support and confidence values.

Association rule is a data mining technique to find association rules or linkages between a combination of items in a set of data. An association rule is said to be interesting, if it meets a minimum value of 2 parameters, namely support and confidence. Support is the percentage of combination items appearing in the database. Confidence is the percentage of strong relationships between items. Examples of association rules:

$$\text{Cappucino} \Rightarrow \text{Donat} [\text{support} = 1\%, \text{confidence} = 50\%]$$

50% confidence means that if a customer buys cappuccino, there is a possibility that 50% of the customers will buy donuts too, or in other words, from 50% of transactions in the database containing Cappuccino items, also contain Donuts items. Whereas support of 1% means that 1% of all transactions examined show that Cappuccino and Donuts are purchased simultaneously.

The purpose of the association rule is to find all association rules that meet the minimum requirements for support (minimum support) and minimum requirements for confidence (minimum confidence). Association rules that meet minimum support and minimum confidence are said to be interesting rules.

Algorithm Apriori, including types of association rules in data mining. A priori algorithm is often used in shopping cart analysis to find out items that are often purchased simultaneously by consumers [2].

A priori algorithm uses a repetitive approach called level-wise search, where n-itemset is used to find (n + 1) -itemset. The first phase of a 1-itemset is searched by tracing data in the database to determine the number of occurrences of each item. Then calculate the support value and collect items that meet the support value. Support values can be obtained by using the following formula:

$$\text{Support (A)} = \frac{\Sigma \text{transactions containing A}}{\text{Total transactions}}$$

Items that meet the specified minimum support are expressed as L1. Furthermore, L1 is used to search for 2-itemset or L2 by doing a join process or a combination of existing itemsets. After L2 is formed, L2 is used to form L3 (3-itemset) and until the process stops when there are no itemset that meets the support value. After finding a collection of itemset that meets the support value, the association rule is established that meets the minimum confidence set, by calculating the confidence value of each itemset, by the formula:

$$\text{Confidence (A} \Rightarrow \text{B)} = P(\text{B}|\text{A}) = \frac{\Sigma \text{ transactions containing A and B}}{\Sigma \text{ transactions containing A}}$$

There are two main processes in a priori algorithm, namely:

- a. Join, This process is done by combining items with other items until no combination can be formed.
- b. Pruning, The pruning process is the result of the items that have been combined and then trimmed using the minimum support specified by the user

3. Methods

This Apriori Algorithm is very suitable for use in purchasing transaction analysis, or also called market basketball analysis to find association rules between purchased items. So the writer uses a priori algorithm as the material of this study. The results of using a priori algorithms are very appropriate as a source of information for problems that exist in the promotion development process, knowing the purchase pattern by consumers, producing the right menu association rules to be used as a decision making [3].

Business Understanding, In this stage is to understand the problem, goals and needs from a business point of view, it takes an understanding of the data mining activities that will be carried out. In this case, the identification of the problems found in the O! Fish restaurant is the determination of a promotional strategy that has not boosted sales, and the procurement of raw products that are difficult to predict. In understanding the problems that exist in the object of research it is very important to be able to determine the initial strategy that is more on target and can achieve goals. Taking new information is a step that can be done in this research activity

Data Understanding, This phase is the collection of initial data that will be used or processed later. The data that will be used is data from consumer purchase transactions stored in a database that occurs within a period of 5 days of O! Fish restaurant operations (April 11, 2018-April 15, 2018). This data will be used for association rules analysis. This data the author limits in its usage, the data used is transaction data that occurs within 5 days, because in the a priori algorithm that is examined in each transaction not how many items each item is sold, but what is examined is whether an item is sold or not.

Data Preparation, From understanding the data in the understanding phase data then the data that has been collected, the author moves the transaction data that has been obtained into the form of Excel file to facilitate data processing, by eliminating some attributes that are not used. Of the many types of food and beverage menus available at O! Fish restaurant, the data taken here only uses 38 types of menus, 23 types of food menus and 15 types of drinks because the authors did the survey, from the number of types of food and drinks, data types These foods and drinks often appear or are often ordered by consumers.

Modelling, The phase where we design a study, conceptualize research starting from the preparation of the data to be used, the selection of data mining techniques to be used, and combining the parameters of the data mining techniques with optimal values. Until the evaluation year and end with the conclusion of the results.

Evaluation, At this stage the prepared dataset will then be processed using the Rapid Miner 8.2 application. This stage starts preparing the sample data that has been determined to be processed and analyzed. The evaluation here will be carried out in depth with the aim of adjusting the results at the modeling stage to suit the objectives to be achieved in the Understanding Business stage [4][5].

1. **Evaluation result.** This stage is the stage where finding association rules that can be used as new sources of information, find the menu association rules from strong transaction data that meets the minimum support and confidence values. Declared a good association rules.
2. **Review Process.** The following is the process for finding menu association rules that meet the two parameters of the support value and the confidence value of the a priori algorithm. Given examples of datasets used in the evaluation process, data processing uses rapid miner 8.2. as follows :

Table 1. Sample Data

	Cobia Pan	Lchee Tea	Fry Dory	Thai Ice Tr	Nasi Soka	Kiwi Rifre	S.Spicy Tu	Smoked S	Mineral W	Black Don	Infused W	Salmon P	Thai Gree	Dory Pan	Ice Lemon	Fettuccini	Olfish Ice	Olfish Gar	Sea Breze	Fry Calam	Green
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	1
4	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
5	0	0	0	0	1	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0
6	1	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0
7	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0
8	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0
9	0	0	0	1	1	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0
10	0	0	1	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	1	0
11	1	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
12	0	0	1	0	0	0	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0
13	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0
14	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
15	0	0	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
17	0	0	0	1	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0
18	1	0	0	0	0	0	1	0	1	0	0	0	0	0	1	0	0	0	1	0	0
19	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	0	0	1	1	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0
21	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
22	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0
23	1	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0

4. Results

A priori algorithm is successfully used to calculate and find association rules with good results, in accordance with the two important parameters in the association rules using a priori algorithms, namely the parameters of minimum support values and minimum confidence values. The level of accuracy of the association rules can be determined by the analyst, if the analyst wants to get association rules with a high degree of accuracy, then a minimum percentage of support and a confidence value with a high percentage can be set. And vice versa if you want to get a lot of association rules but a small degree of accuracy, you can set a minimum support value and a small confidence value.

The result of association rules (menus) is taken from the transaction dataset that is in a database of 150 transaction data. And calculated using rapid minner. In the test this time the author wants to determine the rules of association with a high degree of accuracy that has a strong relationship between items, so here the author sets a minimum support value of 4% and a minimum confidence value of 60%. From processing datasets using rapid minner, get the results of association rules as follows:

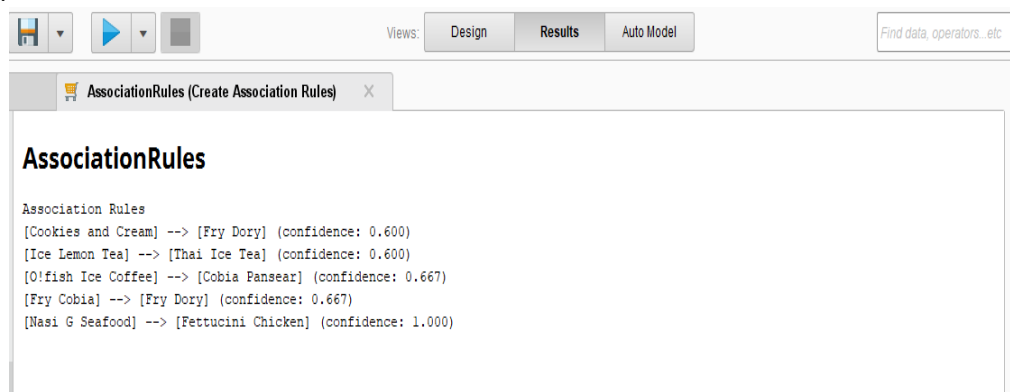


Figure 1 Association Rules (Items) Produced

From the association rules generated 7 association rules 2 the combination of process result itemset based on two parameters of 4% support value and 60% confidence value, support is the percentage of item combinations that appear simultaneously in the dataset. Confidence is a percentage of the strength of relationships between items. The result of the association rule is stated in "if x, then y". Then from the results of one of the rules of association formed in Rapid Minner representing a priori algorithm patterns, for example Premises "O! Fish Ice Coffee Conclusion" Fry Dory "can be stated as follows: If buying" O! Fish Ice Coffee "it will buy Fry Dory "With 4% support and 66.7% confidence.

5. Discussion

The first stage in the search for association rules using a priori algorithm is to determine the minimum value of support and the minimum confidence value to determine how strong the association rules will be sought. Look for frequent itemset that meet the minimum support that has been set. The minimum support value and minimum confidence value are determined by the analyst in this case the author. In determining the minimum support value and the minimum confidence value here is done according to the needs and available data. for example, to find association rules in purchasing items in minimarkets, minimum support is set at 20% and minimum confidence is 70%, but it is different when we try to detect acts of terrorism. Because the data that contains or relates to acts of terrorism is small from the total data available.

The lower the minimum support value and the minimum confidence value, the more association rules are generated, but the lower accuracy values are generated. Because in a matter of a priori algorithm, the rules of a weak purchasing association will also be raised in the results. Conversely, if we give a minimum support value and a high minimum confidence value, the accuracy value produced

will also be good. Because only items that have a strong association in consumer purchases will be displayed in the calculation results.

6. Conclusion

Based on the results of the study, digging up data by utilizing the field of science using data mining to find new information patterns, as well as analysis of the application to know the sales pattern (item association rules) in the O! Fish restaurant that has been submitted in previous chapters, it can be concluded as follows :

1. It is known that the field of data mining using a priori algorithms can produce new information patterns that are useful that can be used to determine sales patterns (item association rules).
2. The application of knowing the sales pattern (item association rule) the implementation of a priori data mining science algorithm provides information on the item association (menu) rules in consumer purchase patterns that can be used by O! Fish restaurants as promotional development strategies based on items (menus) purchased simultaneously to boost sales.
3. This application produces item association rules or menu combinations in consumer purchasing patterns for promotion strategies that are right on target and accurate compared to using promotional strategies manually.

References

- [1] Tampubolon K Saragih H Reza B Epicentrum K Asosiasi A and Apriori A, 2013 Implementasi Data Mining Algoritma Apriori Pada Sistem Persediaan Alat-Alat Kesehatan *Inf. dan Teknol. Ilm.*
- [2] Nofriansyah D, 2014, Konsep Data Mining Vs Sistem Pendukung Keputusan, *Deepublish.* .
- [3] Gunadi G and Sensuse D I, 2012 Penerapan Metode Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Buku Dengan Menggunakan Algoritma Apriori Dan Frequent Pattern Growth (Fp-Growth) : *Telematika.*
- [4] Yanto R and Khoiriah R, 2015 Implementasi Data Mining dengan Metode Algoritma Apriori dalam Menentukan Pola Pembelian Obat *Creat. Inf. Technol. J.*
- [5] Özyer T Erdem Z Rokne J and Khoury S, 2013 *Mining Social Networks and Security Informatics .*