

# ICEBM2015

*by* Viny M

---

**Submission date:** 08-Nov-2020 02:21PM (UTC+0700)

**Submission ID:** 1439394082

**File name:** ICEBM2015.pdf (610.54K)

**Word count:** 2976

**Character count:** 15588

12  
**E-COMMERCE APPLICATION WITH APRIORI METHOD  
TO DETERMINE TOP AND RECOMMENDATION PRODUCT**

Bagus Mulyawan<sup>1)</sup>, Viny Christanti M.<sup>2)</sup>, Riyan Wenas<sup>3)</sup>

16  
<sup>1) 2), 3)</sup> Faculty of Information Technology, 143 Hanagara University  
Jalan S.Parman No.1 Jakarta 11140 Indonesia

<sup>1)</sup> bagus@fii.untar.ac.id, <sup>2)</sup> viny@fii.untar.ac.id, <sup>3)</sup> riyan\_wenas\_94@yahoo.com

**Abstract**

*E-commerce is online media to make selling, purchasing and marketing. There are lots of features in E-commerce that have goal to increase selling. One of them is a feature that displays the items most frequently purchased and items purchased concurrently. "Istana Accessories" Store, is one of the stores that want to use e-commerce to boost sales service to their customers. In this research, we make E-commerce web based application with focus on determining the features of top and recommendation product. Top and recommendation product obtained by doing market basket analysis, so we can get the association rule between items. By display product recommendations, the store has been doing cross-selling process. Apriori method applied to determine top and recommendation product in general based on transaction record. Apriori identifying the frequent individual items and extending to the larger items as long as those items set appear quite often in transaction. The results show that top product and recommendation product from the applied of Apriori algorithm successfully obtained and displayed on the customer page. The test value obtained as the best value with Top minimum support count is 8, Recommendation minimum support count is 3 and Minimum confidence is 50%.*

**Keywords:** E-Commerce, Cross-Selling, Apriori Method, Top Product and Recommendation Product

**Introduction**

E-commerce (electronic commerce or EC) is the buying and selling of goods and services, or the transmitting of funds or data, over an electronic network, primarily the Internet (Turban, 2008). Currently, e-commerce has been very widely used as a medium of sales that are considered to reach various groups and places. It cannot be denied that the items sold on e-commerce have not only the goods with high prices, but also second-hand items sold through e-commerce. Starting from lazada.com, bukalapak.com, tokopedia.com, blibli.com and many more forms of e-commerce that began to appear in the virtual world.

Forms of e-commerce are presented unique. But many e-commerce only display photos of the goods and only equipped with a category of goods. Displaying the goods as it is, sometimes make it difficult for buyers to select or search for items according with their wishes. One way to do marketing and increase sales is by creating or displaying the goods that are interconnected. Arrange the goods regularly facilitate the buyer to see the linkage the goods with each other. If the e-commerce has been able to show the items concerned, then the next stage is to give discounts if the goods are purchased simultaneously.

The process of searching for the relationship between the objects with each other is not easy. One way is to record every transaction data that has happened then analyzed the relationship between one items to another. Results of this analysis can help the seller to determine the most frequently purchased item, goods that can be a recommendation because it is related to each other.

Currently, many studies were done to make the process of analysis of sales data automatically. The analysis process can automatically simplify and accelerate the application in determining the best-selling products and products that are interrelated. The process of recording and analyzing the sales data can be performed by the method of data mining.

Based on 15 background, in this study the analysis of the customer's buying process is carried out using association rules in data mining. Association rules can also be called a market basket analysis, where the analysis is performed against the purchase transaction has occurred. The analysis process should be done with a lot of data transactions in a given period. Market basket analysis was conducted to determine the relationship between products based on purchases already made. If the analysis of the product purchase has been made, the results of this analysis can be used to develop a marketing strategy.

5 One of the methods used in the search for association rules on transaction data is to use Apriori algorithm. The Apriori algorithm is a classical data mining method for association rule discovery typically applied to market basket data, such as the study of what products tend to be purchased together in an on-line market place (Maheswari, 2014). There are two approach in this algorithm, bottom up and top-down approach. The classical Apriori algorithm follows bottom up approach. In this research, we use top down approach, where in the rules are generated by avoiding generation of unnecessary patterns. The major advantage of this approach is that, the number of database scan and times are reduced.

In this study, the data used is the data at the “Istana Aksesoris” Store. “Istana Aksesoris” store is a store that sells mobile phone accessories and equipment. These stores can sell goods wholesale or retail. Most sales are sales made in the form of wholesale. During this time “Istana Aksesoris” doing sales as manually, where buyers place an order in person or by telephone. The shop is hoping to increase the number of sales and increase the number of customer. Therefore, these results are applied in the form of the creation of e-commerce for the “Istana Aksesoris”. E-commerce made equipped with features that can show top and recommendations product based on the analysis of transaction data that is done automatically.

#### E-commerce

E-Commerce or so-called electronic commerce is spread, purchasing, sales, and marketing of goods via the Internet or computer network (Laudon and Trevor, 2009). The advantages of the use of e-commerce that is able to overcome the problem of integration, publication, interaction and transaction. Typically, the menu contained in e-commerce is a shopping cart, payment methods, contact person and the search menu. But further, web-based e-commerce features that support aimed at increasing sales. Such features include a menu that can display product category, latest products, special offer, payment, sitemap, testimonials, top product, recommendation product, and related product.

Some important features to be presented are the top product, recommendation product, and related product. The top product menu is displayed with the aim that the customer can find out information of the most frequently purchased product (top product). While the recommendation product menu is displayed for expected that customer will add the type of goods that relevant with the first product (recommendation product). And the related product is references option for other similar products with different brands (related product).

Analysis of the association or the association rule mining is data mining techniques to discover the rules of associative between combinations of items (Han and Kamber, 2006). Examples of association rules of analysis of purchases in a supermarket are able to know how likely a customer to buy bread along with milk. With the knowledge of the owner of the supermarket can adjust the placement of the goods or designing a marketing campaign by using a combination of discount coupons for certain items. The association analysis become famous because of its application to analyze the contents of a shopping cart in the supermarket, association analysis is also often referred to as market basket analysis (Linoff and Berry, 2011).

#### Apriori Algorithm with Top-Down Approach

Apriori is designed to operate on databases containing transactions (for example, collections of items bought by customers, or details of a website frequentation (Agrawal and Srikant, 1994). There are two approach of Apriori algorithm, bottom-up and top-down approach. Apriori algorithm with a top-down approach chosen because it has advantages, the process does not generate all frequent itemset like the classic Apriori thus reducing unneeded pattern itemset (Maheshwari, 2014). This is expected to create a data mining time becomes shorter.

Apriori is a method that follows the analysis of the association which is trying to find a technique of associative rule between combinations of items. Such as assessing, how frequently purchased item A together with item B. It is important or not of association rule, assessed by two parameters, namely the support and confidence. Support is the percentage of the combination of the items in a set of items in the transaction. While confidence is a strong relationship between the values of items in the collection of the item.

For example, if known {hand phone, case} → {screen protector} gets value of support is 40% and confidence is 50%. Could mean that, handphone, case and screen protector purchased together occurs with probability 50%. The transaction is valid because 40% of people of all transactions making a purchase like that.

Apriori is a classic algorithm for learning association rules in data mining. This algorithm is an iterative approach and there are two steps on each iteration. The first step generates a set of candidate itemsets. Then, in the second step we count the occurrence of each candidate set in database and prune all disqualified candidates (i.e. all infrequent itemsets). The algorithm is based on the closure property of frequent itemsets: if a set of items is frequent, then all its proper subsets are also frequent. There are two properties (Maheshwari, 2014):

1. Generate all frequent itemsets: One frequent itemsets is an itemsets which support higher than minimum support (min\_sup).

$$support(A \cap B) = \frac{\text{transaction contains A and B}}{\text{total transaction}} \dots(1)$$

2. Generate all confident association rules from frequent itemsets: One confident association rule is a rule which confidence level greater than minimum confidence (min\_conf).

$$confidence = p(A|B) = \frac{\text{transaction contains A and B}}{\text{transaction contain A}} \dots(2)$$

Flowchart of the Apriori algorithm with Top-Down approach can be seen in Figure 1.

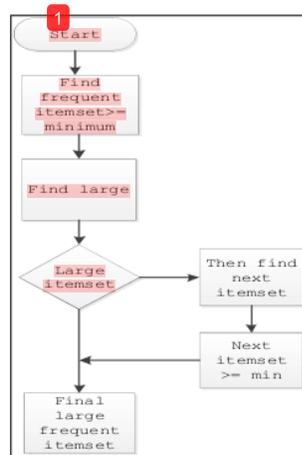


Fig.1 Flowchart Algoritma Top-Down Apriori Method

At Apriori algorithm, the top product obtained by the largest itemset candidates who meet the minimum support count so made as frequent itemset. Item that is in frequent itemset is shown as a top product. While the recommended products are products that are based on the candidate itemsets containing products related to the selected product. Candidates who fulfill used as frequent itemset and value his support and confidence. If it meets the specified value of the items in a frequent itemset can be displayed as recommendation product.

**Result and Discussion**

Based on a survey, it is known that this store has a lot of salesmen in charge of promotion and send the goods to the customer. The process is analyzed to see how the real store policy to determine the top products and product recommendation for promotion to customers. According to the shop owner, a product said to be a top product in terms of the much sought after or in the sense that many appear in the sales transaction. Further to recommendation product, obtained by the frequency of an item purchased together with other goods.

E-commerce websites are created using the programming language PHP. The software used to design the look and build a web site is Adobe Photoshop CS6 and Adobe Dreamweaver CS6. MySQL is used for building and managing databases. While Apache is choose as web server for running PHP. This website is built in two parts, to customers (front end) and store (back end). E-commerce website has 8 main features, namely, Home, Product, Chart, Retur, Customer Profile, Payment Method, Shipping Information and Contact Person. At Figure 2 we can see the interface of “Istana Aksesoris” e-commerce. We can see the most wanted product at the Home menu and we can see the recommendation product for selected product at the Product menu.



Fig.2 The interface of “Istana Aksesoris” e-commerce

Testing top product and recommendation product from the application of Apriori algorithm done using real data from the store. The data used is the sales transaction data, products and customers. The test is divided into two conditions, namely the use of data 3 months and 6 months. 3 months of data taken from January 1, 2014 until March 30, 2014, consisted of 133 transactions. Data 6 months from January 1, 2014 until June 30, 2014, consists of 327 transactions.

Based on the experimental results, it is known that the data of 3 months with top minimum support count is 9 only produce 2 products as the top product. It is considered that minimum support count produces too little number of top products. Then the data of 6 months to experiment with the recommendation minimum support count is 3 and minimum confidence is 75% obtained recommendation product only 4. It was thought a product that has too little product recommendation. After testing, the test value obtained as the best value is as follows: Top minimum support count is 8, Recommendation minimum support count is 3 and Minimum confidence is 50%. At table 1 we can see the top product result and at table 2 we can see the recommendation product result that using that score.

**Table 1** Top Product with top minimum support count = 8

Period	Date		Top Product
3 month	January 1, 2014	March 30, 2014	VIORA 5600MAH VIORA 8400MAH FUZE IPHONE 4 FUZE IPHONE 5
6 month	January 1, 2014	June 30, 2014	FLIP CASE GALAXY ACE 3 FLIP CASE GALAXY CORE FLIP CASE GALAXY GRAND FLIP CASE GALAXY S4 VIORA 5600MAH VIORA 8400MAH FUZE IPHONE 4 FUZE IPHONE 5

**Table 2.** Recommendation Product with minimum support count = 3 and minimum confidence = 50%.

Period	Date		Product Choice	Recommendation Product
3 month	January 1, 2014	March 30, 2014	FLIP CASE UNIVERSAL 7" PICT	FLIP CASE UNIVERSAL 7"
			FLIP CASE GALAXY S4	FLIP CASE GALAXY CORE
			FLIP CASE LENOVO A516	FLIP CASE LENOVO A369I
			BATERAI HIPPO BB 9900	BATERAI HIPPO BB 9700
			BATERAI HIPPO BB 9700	BATERAI HIPPO BB 9900
			BATERAI BB 9700 ORI	BATERAI BB 9800 ORI
			BATERAI BB 9300 ORI	BATERAI BB 9300 ORI
			BATERAI BB 9800 ORI	BATERAI BB 9700 ORI
			VIORA 5600MAH	VIORA 8400MAH
			VIORA 8400MAH	VIORA 5600MAH
			FUZE IPHONE 5	FUZE IPHONE 4
			FUZE IPHONE 4	FUZE IPHONE 5
6 month	January 1, 2014	June 30, 2014	FLIP CASE UNIVERSAL 7" PICT	FLIP CASE UNIVERSAL 7"
			FLIP CASE GALAXY ACE 3	FLIP CASE GALAXY CORE
			FLIP CASE GALAXY S4	FLIP CASE GALAXY CORE FLIP CASE GALAXY GRAND
			FLIP CASE ANDROMAX I2	FLIP CASE ANDROMAX C
			BATERAI HIPPO BB 9900	BATERAI HIPPO BB 9700
			BATERAI HIPPO BB 9700	BATERAI HIPPO BB 9900
			BATERAI BB 9320 ORI	BATERAI BB 9300 ORI
			BATERAI BB 9900 ORI	BATERAI BB 9300 ORI
			BATERAI BB 9800 ORI	BATERAI BB 9700 ORI
			VIORA 5600MAH	VIORA 8400MAH
			VIORA 8400MAH	VIORA 5600MAH
			FUZE IPHONE 5	FUZE IPHONE 4
FUZE IPHONE 4	FUZE IPHONE 5			

The test results of top product of the table 1 shows that for three months of data, acquired four products which become the top product is VIORA 5600mAh, VIORA 8400MAH, FUZE IPHONE 4, IPHONE FUZE 5. This means that within a span of 3 months (January 1, 2014 to 30 March 2014), many customers buy 4 of these products. 5600mAh VIORA products are purchased simultaneously with 8400MAH VIORA and FUZE IPHONE 4 lots purchased together with FUZE IPHONE 5.

Results of testing the top product for data of 6 months, gained 8 products become the top product is FLIP CASE GALAXY ACE 3, FLIP CASE GALAXY CORE, FLIP CASE GALAXY GRAND, FLIP CASE GALAXY S4, VIORA 5600mAh, VIORA 8400MAH, FUZE IPHONE 4, FUZE IPHONE 5. This means that within a span of 6 months, many customers buy 8 products. FLIP CASE product GALAXY ACE 3 lots purchased together with the FLIP CASE CORE GALAXY, GALAXY GRAND CASE FLIP lot purchased together with the FLIP CASE GALAXY S4, VIORA 5600mAh lot purchased together with VIORA 8400MAH, and FUZE IPHONE 4 lots purchased together with FUZE IPHONE 5.

### **Conclusion**

The results show that top product and recommendation product from the applied of Apriori algorithm successfully obtained and displayed on the customer page. The test value obtained as the best value with Top minimum support count is 8, Recommendation minimum support count is 3 and Minimum confidence is 50%. Support count value may affect the amount of product produced. When value is used as the bare minimum is getting smaller, the less the amount of goods that can be obtained. However, if the minimum value is made smaller, it will affect the relationship between products. The smaller the minimum value linkages between product confidences mean getting away. Besides the characteristics of the transaction data also affect the results of the analysis so that the experiment must be carried out by using different values.

E-commerce websites are expected to increase sales of Istana Accessories store. Furthermore, the website is still to be developed in order to have features that maximum. For example personalize based on customer characteristics or add period time transaction data that can be further evaluated. Besides online payment feature and other features can be added to support this website.

### **References**

- Turban, E., Leidner, D., McLean, E., and Wetherbe, J. (2008). *INFORMATION TECHNOLOGY FOR MANAGEMENT, (With CD)*. John Wiley & Sons.
- Maheshwari, Shaikha and Jain, Poja (2014). "The Research on Top Down Apriori using Association Rule". *International Journal of Advanced Research in Computer Science and Software Engineering*, Vol. 4, Issue 4.
- Laudon, K. C., and Traver, C. G. (2009). *E-commerce. Business, Technology, Society*. Editor Pearson, New York.
- Han, J., Kamber, M., and Pei, J. (2011). *Data mining: concepts and techniques: concepts and techniques*. Elsevier.
- Linoff, Gordon S. and Berry, Michael J. A. *Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management*. 3<sup>rd</sup> Edition. New York: Wiley, 2011.
- Agrawal, R., and Srikant, R. (1994, September). Fast algorithms for mining association rules. In Proc. 20th int. conf. very large data bases, VLDB (Vol. 1215, pp. 487-499).

## ORIGINALITY REPORT

18%

SIMILARITY INDEX

17%

INTERNET SOURCES

11%

PUBLICATIONS

13%

STUDENT PAPERS

## PRIMARY SOURCES

1	<a href="http://ijarcsse.com">ijarcsse.com</a> Internet Source	4%
2	Submitted to Udayana University Student Paper	2%
3	Submitted to Swiss German University Student Paper	2%
4	<a href="http://ijircce.com">ijircce.com</a> Internet Source	1%
5	Submitted to Indian Institute of Management, Indore Student Paper	1%
6	Submitted to Universitas Dian Nuswantoro Student Paper	1%
7	<a href="http://henrybusinessblog.blogspot.com">henrybusinessblog.blogspot.com</a> Internet Source	1%
8	Alexander Setiawan, Gregorius Satia Budhi, Djoni Haryadi Setiabudi, Ricky Djunaidy. "Data Mining Applications for Sales Information System Using Market Basket Analysis on	1%

# Stationery Company", 2017 International Conference on Soft Computing, Intelligent System and Information Technology (ICSIT), 2017

Publication

---

9	<a href="http://www.cs.cuhk.hk">www.cs.cuhk.hk</a> Internet Source	1%
10	Submitted to American University of the Middle East Student Paper	1%
11	Submitted to Universitas Amikom Student Paper	1%
12	<a href="http://repository.widyatama.ac.id">repository.widyatama.ac.id</a> Internet Source	1%
13	Dedi Trisnawarman, Tris Sutrisno, Novario Jaya Perdana, Sugesti. "Bayesian Networks Modeling Using Partial Least Squares Approach to Predict Stroke Disease", IOP Conference Series: Materials Science and Engineering, 2019 Publication	<1%
14	<a href="http://co.ontario.ny.us">co.ontario.ny.us</a> Internet Source	<1%
15	"Web Data Mining", Springer Science and Business Media LLC, 2007 Publication	<1%

---

16

Teny Handhayani, Ito Wasito. "Fully unsupervised clustering in nonlinearly separable data using intelligent Kernel K-Means", 2014 International Conference on Advanced Computer Science and Information System, 2014

Publication

<1%

---

Exclude quotes On

Exclude matches Off

Exclude bibliography On