RESEARCH ARTICLE OPEN ACCESS

Association Algorithm Modeling for Sales of Steel-Manufactured Products in Indonesia

Wendi Usino*, Arief Wibowo**
*Universitas Budi Luhur, Indonesia
Email: wendi.usino@budiluhur.ac.id
**Universitas Budi Luhur, Indonesia
Email: arief.wibowo@budiluhur.ac.id

----------**********----------

Abstract:

As the fourth largest country in the world, Indonesia has several manufacturing companies that produce steel. The challenges these companies face include changes in foreign exchange rates, such as the dollar against the rupiah, and competitive business competition due to many imported products on the steel metal market. On the other hand, the vast potential for the steel product market forces national companies to have an adaptive sales strategy by adjusting the sales pattern to have maximum revenue. A business intelligence approach is needed to determine how a product item is associated with other known things from the purchase transaction database to maintain the sales value of steel products for manufacturing companies. This study proposes a data mining model that applies the association method to determine the relationship patterns and relationships between variants of steel products purchased by consumers. By knowing the association's relationship, new marketing patterns can be offered to increase sales value. This study compares two types of association algorithms, namely Apriori and Fp-Growth. The results showed that the a priori algorithm is the best model for the cross-selling scheme of steel products from manufacturing companies in Indonesia. The association rules are obtained, which can be implemented in the cross-selling strategy.

Keywords — **Apriori Algorithm; Intelligence Bussiness; Data Mining; Sales Patterns**.

----------**********----------

I. INTRODUCTION

In the era of free trade, the presence of imported products is a challenge faced by various companies. In Indonesia, steel manufacturing companies are one of the industry players who must face these challenges. One company's downfall proves the difficulties of steel manufacturing companies in Indonesia after another. Even during the last five years, Indonesia's national steel manufacturing company has faced dire conditions. The tight business competition requires companies to have the ability to manage business intelligence so that they can be a resource in determining strategies to maintain and win business. One of the things a steel company can do to gain savvy business knowledge is to analyze purchasing patterns as a strategic resource for product marketing. The data mining approach is one of the best models that can be used to gain valuable business knowledge. One well-known method is the application of an a priori algorithm that can find associations between product items in past purchase data.

This study analyzes purchase data with one algorithm and compares it with a comparison algorithm. Thus, analyzing the two methods will be one of the more optimal ways for companies to gain knowledge and implement sales strategies for steel products in Indonesia.

This study aims to determine a sales strategy with maximum revenue prediction from modeling using an a priori data mining algorithm and FP-growth comparisons. This study will analyze the sales transaction data of steel material products and their components with a market basket analysis method to find several association rules of sales transactions. The Apriori algorithm and the Frequent Pattern Growth (FP-Growth) algorithm are used. These two algorithms are used to form frequent item sets, which will later be used as a reference for formulating the association rules for selling steel component materials. The association rules generated by the Apriori algorithm and FP-Growth will be evaluated and analyzed further to predict the value of future sales. The best association rules can be used to determine the cross-selling marketing strategy and product bundling marketing techniques. The formulation of the problem that is solved in this study is what type of association algorithm is best able to generate knowledge about the purchasing patterns of steel products based on the evaluation indicators of association rules.

II. RESEARCH METHODOLOGY

This study adopts a data mining model to obtain business knowledge from purchasing data analysis. The model adopted from previous studies uses the Knowledge Discovery in Database (KDD) approach, as shown in Figure 1.

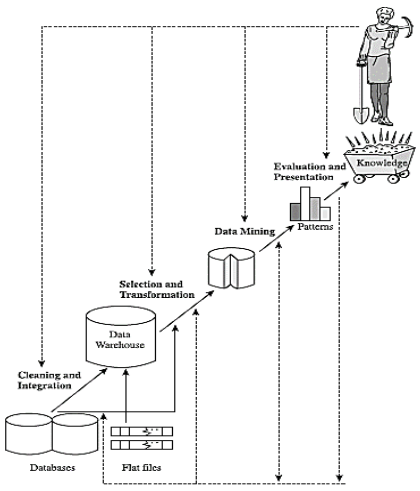

Fig. 1 The Knowledge Discovery in Database Concept

Based on the KDD approach, the discovery of business knowledge can be described in the following processes:

1) Data Cleaning
   At this stage, selecting relevant data from the database is done by separating inconsistent data and irrelevant data.
2) Data Integration
   At this stage, integration of existing data is carried out by combining various data sources into one source.
3) Data Selection
   At this stage, the selection of data relevant to the analysis will be carried out in the database.
4) Data Transformation
   At this stage, changes are made to the existing data format into a suitable data format for processing in data mining.
5) Data Mining
   At this stage, the data mining process is carried out by applying specific methods to obtain confidential information from existing data.
6) Pattern Evaluation
   The identification of exciting patterns obtained from data mining is carried out and then represented at this stage.
7) Knowledge Presentation
   At this stage, visualization and presentation of knowledge about the techniques used to obtain user knowledge are carried out.

A business analysis approach called Market Basket Analysis is needed, which is a methodology for analyzing consumer buying habits by finding associations between several different items, which consumers place in a shopping basket purchased in a particular transaction. The purpose of market basket analysis is to determine which products may be bought simultaneously (Han and Kamber, 2006).

The term market basket analysis comes from a common practice in supermarkets: when consumers put all the goods, they buy into a basket (basket), which is generally provided by the shop. Information about products that consumers usually buy together can offer insights to shop or

supermarket managers to increase their business profits.

Association rule mining is a data mining technique to find association rules between a combination of items (Kusrini, 2009). The impressive measure that can be used in data mining is Support, which is a measure that shows the level of dominance of an item or itemset from the entire transaction. In addition, there is also another parameter, namely Confidence, which is a measure that shows the relationship between two items in a conditional (based on a particular condition).

The association analysis methodology is divided into two stages: analyzing high-frequency patterns (frequent patterns) and forming association rules. Association rule is a method commonly used to find relationships between items. For example, from a set of transaction data, a person finds a relationship like the following. When a customer buys a laptop, he usually also buys a mouse in the same transaction. When a customer buys a toothbrush, he also buys toothpaste. Thus, finding the relationship between these items may require reading transaction data repeatedly in a large amount of transaction data to find different relationship patterns, so the computation time and costs, of course, will also be considered to see the relationship. It requires an algorithm that has high accuracy.

The a priori algorithm is a basic algorithm proposed by Agrawal & Sikant in 1994 to determine frequent itemsets for the Boolean association rule (Han and Kamber, 2000). The priority algorithm uses knowledge of previously known attribute frequencies for further information processing. The a priori algorithm determines possible candidates by paying attention to minimum Support and minimum Confidence. Support is the visitor value or the percentage of a combination of items in the database. The support formula is as follows:

$$Support\ (A) = \frac{Number\ of\ Transactions\ containing\ A}{Total\ Transactions}$$

The two-item Support value is obtained using the formula:

$$Support(A, B) = P(A \cap B)$$

$$Support(A, B) = \frac{\sum Transactions\ containing\ A\ and\ B}{\sum Transaction}$$

While confidence is the value of certainty, which is the strength of the relationship between items in an a priori, confidence can be searched after the frequency pattern of an item's appearance is found. The formula for calculating confidence is as follows:

For example, the rule A → B is found, so:

$$Confidence\ P(B|A) = \frac{Total\ Transactions\ containing\ A\ and\ B}{Transactions\ containing\ A}$$

The primary process carried out in the a priori algorithm to get frequent itemset is Join. This process is done by combining items with other items so that they cannot be combined. The next stage is Prune (Pruning), which results from the combined items then trimmed using the minimum support set by the user. The principle of the a priori algorithm includes collecting single items, then looking for the largest item, then getting candidate pairs then calculating the large pair of each item. The next step is to determine the candidate triplets of each item and so on. Each subset of a frequent itemset must be regular. The form of the a priori algorithm can be written as follows:

```
L1 = {Frequent itemset with one element};
for ( k=2; Lk-1 ≠0; k++)
{
        Ck = apriori-gen(Lk-1); // New candidates
        For all transactions t
        {
                Ct = subset(Ck,t); // Candidates
                contained in t
                for all candidates c ∈Ct do
                c.
        count++;
        }
        Lk = {c ∈Ck | c.count >= minsup }
        }
return = Uk Lk ;
```

Where :

L : set of frequent itemset
minsup: minimum support
C : the set of candidate itemset
c : itemet candidate
t : transaction

The FP-Growth (Frequent Pattern-Growth) algorithm is an alternative to finding frequent itemset without using candidate generation (Wu and Kumar, 2009). In determining the frequent itemset, there are two process stages that must be carried out, namely the creation of an FP-tree and the application of the FP-Growth algorithm to find frequent itemsets. The data structure used to search for frequent itemsets with the FP-Growth algorithm is an extension of the use of a prefix tree, which is commonly called the FP-tree, the FP-Growth algorithm can directly extract frequent itemset from the FP-tree that has been formed using the divide and conquer principle.

A. Association Rule Evaluation

There are 2 factors that will be evaluated in this study, namely the measure of generality and the measure of reliability of the resulting association rules. The measure of generality is used to determine the level of appearance of each item formulated in the association rules for the entire transaction. The sizes used are:

1. Support, which is a measure of how often a collection of items in association occurs together as a percentage of all transactions, is similar:

$$s(A \Rightarrow B) = P(A \cap B)$$

with terms and conditions

$$(A \cap B) = \frac{\sum Transactions\ containing\ A\ and\ B}{\sum Transaction}$$

The entire document should be in Times New Roman or Times font. Type 3 fonts must not be

2. *Coverage*, is a measure of how often the collection of each item defined on the left (causal factor) occurs as a percentage of all transactions, with the equation:

$$= \frac{\sum Transactions\ containing\ A}{\sum Transaction}$$

The literature regarding the analysis of association rules used in this study is described in Table I.

TABLE I PREVIOUS STUDIES

No.	Researcher	Method	Research result
1.	Venkatachari, 2016	FP-Growth and Apriori Algorithm	The Apriori algorithm that is applied to the R Programming application, in the first experiment with 0.05 support it takes 1 millisec and produces 111 association rules. In the second experiment with 0.009 support it takes 1 second and produces 2,090,141 association rules, in the third experiment with 0.01 support it takes 2 seconds and produces association rules of 2,090,141. Then the Fp-Growth algorithm is applied in the Rapidminer application, in the first experiment it was carried out with 0.9 support and took 1 second, resulting in 987 association rules. In the second experiment with a minimum support of 0.05 and the specified time is out of order.
2.	Loraine Charlet Annie M.C and Ashok Kumar D, 2012	Apriori and K-Apriori Algorithm	The first experiment was carried out with a dataset size 13468X302, the algorithm aproiori produces 32 rules and K-apriori produces 19532 rules. The second experiment was carried out with a dataset size 9620X302, the algorithm aproiori produced 92 rules and K-apriori produced 2009 rules. The second experiment was conducted with a dataset size 6438X302, the algorithm aproiori yielded 27 rules and K-apriori produced 11731 rules..
3.	Nadya Rahmawati, 2017	Aprori Algorithm	Minimum support 10.86 and Minimum confidence 40.85 yield 9 association rules, minimum support 2.82 and Minimum confidence 40.24 yield 4 association rules

No.	Researcher	Method	Research result
4.	Dyah Pramesthi Larasati, 2015	Fp-Growth Algorithm	With a minimum support of 0.002 and a minimum confidence of 0.5 in January 2014 it produced 2 rules, February 2014 produced 2 rules, March 2014 produced 2 rules, April 2014 produced 2 rules, and in May 2014 produced 1 rule.
5.	Alkadri Masnur, 2015	Aprori Algorithm	Minimum support 0.25 (25%) and Minimum cnfidece 0.75 (75%) results in 3 association rules
6.	Listriani, Setyaningrum and A, 2016	Aprori Algorithm	The 5% minimum support and 15% minimum confidence value results in 7 association rules
7.	Nur Rohman Ardani, 2016	Fp-Growth Algorithm	The most ideal minimum support value is 0.15 and a minimum confidence value of more than 0.6 produces 48 association rules

Thus, this study will use a priori and fp-growth algorithms to determine the best algorithm based on a higher level of accuracy for determining the association of steel products.

III. RESULTS AND DISCUSSION

The data mining process with the apriori algorithm in this study will be carried out on table_of_sales. This table contains sales data with the highest 20% appearance and exceeds or is equal to the minimum limit (minimum support) of 2%. This data is applied to the Rapidminer version 08 application. The transaction data used in training data (training data) are 924 transactions with transaction details 2369. The next step is to create a binary table for each item combination of the sales transaction set. An example of Binary Tables of Apriori Algorithm Training Data is shown in Figure 2.

OrderNumber	JIS G 3101 SS400 -	JIS G 3101 SS400 30	JIS G 3131 SPHC -	JIS G 3131 SPHC 24
6000031617	0	1	0	0
6000031622	0	0	0	0
6000031672	0	1	0	0
6000031677	0	0	0	0
6000031682	1	0	1	0
6000031683	0	0	1	0
6000031686	0	0	1	0
6000031687	0	0	1	0
6000031693	0	0	0	0
6000031700	0	0	0	0
6000031703	0	1	0	0
6000031707	0	1	0	0
6000031710	0	0	0	0
6000031756	0	0	0	0
6000031757	0	0	0	0
6000031759	0	0	0	0
6000031762	0	0	0	0
6000031764	0	0	0	0
6000031768	0	0	0	0
6000031772	0	0	0	0
6000031773	0	0	0	0
6000031782	0	0	0	0
6000031783	0	0	0	0
6000031784	0	0	0	0
6000031786	0	0	0	0

Fig. 2 The Binary of Apriori Algorithm Data Training

The application of data mining with the association method in Rapidminer can be made by forming a binary data table combination of an itemset in formal .xls. Then run the model as shown below:

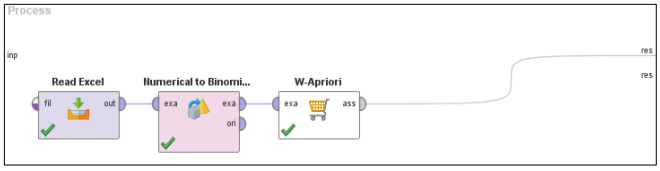

Fig. 3 The Binary of Apriori Algorithm Data Training

The association rules obtained from the application of market basket analysis with the a priori algorithm are as follows:

TABLE II APRIORI ALGORITHM ASSOCIATION RULES IN DATA TRAINING

No.	A		B	Conf.
1	SNI 07 0601 BJPC	==>	JIS G 3131 SPHC	0.99
2	SNI 07 0601 BJPC 30	==>	JIS G 3131 SPHC 30	0.94
3	JIS G 3131 SPHC 30	==>	SNI 07 0601 BJPC 30	0.86
4	JIS G 3131 SPHC	==>	SNI 07 0601 BJPC	0.63

Evaluation of association rules is done by measuring the generality and reliability of each of the resulting regulations. The measure of conception is used to determine each item's level of appearance formulated in the association rules for the entire transaction. At the same time, the

standard of reliability is used to determine the level of reliability of the association rules generated in association rules mining. The evaluation of the Apriori algorithm training data association rules can be seen in the following table:

TABLE III EVALUATION OF THE ASSOCIATION RULES

Sequential	Generalities		Reliability		
	Support	Coverage	Confidence	Added Value	Correlation
1	0.10065	0.10173	0.98936	0.88871	0.09958
2	0.41126	0.43831	0.93827	0.52702	0.38587
3	0.41126	0.47835	0.85973	0.44847	0.35357
4	0.09957	0.15909	0.62585	0.52520	0.06299
Average	0.25568	0.29437	0.85330	0.59735	0.22550

In determining the frequent itemset in the FP-Growth algorithm, two process steps must be carried out, namely creating the FP-tree and applying the FP-Growth algorithm to find frequent itemset. The data structure used to search for frequent itemsets with the FP-Growth algorithm is an extension of a prefix tree, which is commonly called the FP-tree. The FP-Growth algorithm can directly extract frequent itemset from the FP-tree that has been formed. In FP-Tree development, two database searches are required. The first search is used to calculate each item's support value and select products that meet the minimum support value, then sort by the largest number of frequencies or occurrences. In the Rapidminer application, to find out the association rules for the fp-growth algorithm, it can be done by creating binary data from an item set and then saving it in xls format. The next step is to apply this data to the Rapidminer version 08 application with .xls file format and a model like the following:

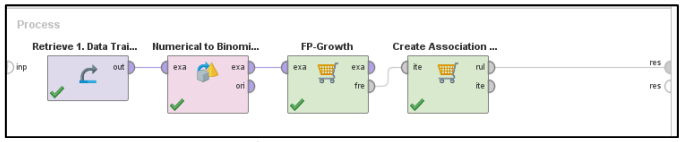

Fig. 4 The Modelling of the Fp-Growth Algorithm at Rapidminer

From this process produces the following association rules:

TABLE IV FP-GROWTH ALGORITHM DATA TRAINING ASSOCIATION RULES

Seq.	A		B	Conf
1	SWRM8 9,0 mm	==>	SWRM8 5,5 mm, SWRM8 10,0 mm	0.897
2	SWRM8 5,5 mm	==>	SWRM8 9,0 mm, SWRM8 10,0 mm	0.907
3	SWRM8 10,0 mm	==>	SWRM8 10,0 mm	0.913
4	SWRM8 5,5 mm	==>	SNI 07 0601 BJPC	0.930
5	SWRM8 10,0 mm	==>	SWRM8 5,5 mm	0.937
6	SWRM8 9,0 mm	==>	SWRM8 5,5 mm	0.944
7	SWRM8 9,0 mm, SWRM8 10,0 mm	==>	SWRM8 5,5 mm	0.947
8	SWRM8 9,0 mm	==>	SWRM8 10,0 mm	0.947
9	SWRM8 9,0 mm, SWRM8 5,5 mm	==>	SWRM8 10,0 mm	0.950
10	SWRM8 5,5 mm	==>	SWRM8 9,0 mm	0.954
11	SWRM8 10,0 mm	==>	SWRM8 9,0 mm	0.964
12	SWRM8 5,5 mm, SWRM8 10,0 mm	==>	SWRM8 9,0 mm	0.975

Evaluation of association rules in the fp-growth algorithm is carried out by measuring the generality and reliability of each of the resulting rules:

TABLE V. EVALUATION OF DATA TRAINING ASSOCIATION RULES

Seq.	Generalities		Reliability		
	Support	Coverage	Confidence	Added Value	Correlation
1	0.00090	0.00100	0.89700	0.89608	0.00083
2	0.00090	0.00099	0.90700	0.90605	0.00086
3	0.00090	0.00098	0.91300	0.91206	0.00086
4	0.00092	0.00099	0.93000	0.92902	0.00091
5	0.00092	0.00098	0.93700	0.93601	0.00093
6	0.00094	0.00100	0.94400	0.94301	0.00093
7	0.00090	0.00095	0.94700	0.94601	0.00094
8	0.00095	0.00100	0.94700	0.94602	0.00093
9	0.00090	0.00094	0.95000	0.94902	0.00093
10	0.00094	0.00099	0.95400	0.95300	0.00095
11	0.00095	0.00098	0.96400	0.96300	0.00096
12	0.00090	0.00092	0.97500	0.97400	0.00098
Rata-Rata	0.00092	0.00098	0.93875	0.93777	0.00092

B. Analysis of Evaluation Results

The analysis of the evaluation results is done by calculating the level of strength of the association rules generated by each algorithm. The association rule's level of strength is determined by the support value, which represents the generality aspect, and the confidence value, which means the reliability aspect. The following is a measure of the level of strength against the resulting association rules:

TABLE VI. COMPARISON OF APRIORI ALGORITHM AND FP-GROWTH RESULTS

Apriori			
No	Supp	Conf	Supp * Conf
1	0.10065	0.98936	0.09958
2	0.41126	0.93827	0.38587
3	0.41126	0.85973	0.35357
4	0.09957	0.62585	0.06231
		Total	**0.22533**

Fp-Growth			
No.	Supp	Conf	Supp*Conf
1	0.00090	0.89700	0.00080
2	0.00090	0.90700	0.00081
3	0.00090	0.91300	0.00082
4	0.00092	0.93000	0.00086
5	0.00092	0.93700	0.00086
6	0.00094	0.94400	0.00089
7	0.00090	0.94700	0.00085
8	0.00095	0.94700	0.00090
9	0.00090	0.95000	0.00085
10	0.00094	0.95400	0.00090
11	0.00095	0.96400	0.00091
12	0.00090	0.97500	0.00087
		Total	**0.00086**

The accuracy rate of the a priori algorithm association rules and fp-growth is measured by calculating the support times the confidence of each resulting association rule. From Table VI it can be seen that the accuracy of the a priori algorithm is 22.53% while the accuracy of the fp-growth algorithm is 0.09%.

Based on the results of the evaluation of the two tested algorithms, it can be concluded that the association rules generated by the Apriori algorithm have better performance than the rules generated from the Fp-growth algorithm. Thus, the a priori algorithm will be implemented into the prototype of the customer shopping pattern analysis system, and the resulting association rules will be used as a reference for the product to be promoted. Based on the formed evaluation results, business knowledge is obtained that can be applied to manufacturing companies in Indonesia in selling steel products. The association rules that are included are as follows:

TABLE VII. PATTERN KNOWLEDGE OF THE ALGORITHM

Association Rules			Knowledges
SNI 07 0601 BJPC	==>	JIS G 3131 SPHC	Pile type steel products, should be sold together with foundation construction steel products in one package
SNI 07 0601 BJPC 30	==>	JIS G 3131 SPHC 30	Steel products for small road construction, should be sold in conjunction with small size drainage steel constructions
JIS G 3131 SPHC 30	==>	SNI 07 0601 BJPC 30	Small size drainage construction steel products can be sold together with small size pile types
JIS G 3131 SPHC	==>	SNI 07 0601 BJPC	Large volume drainage type steel products can be sold together with large volume main piles

IV. CONCLUSIONS

From all stages of the research that has been carried out, several conclusions can be drawn as follows:

1. The Apriori Algorithm can be used as a method to find the best association rules in data analysis of steel product purchases in Indonesia.
2. The Apriori algorithm has a better performance than the FP-growth algorithm, it can be seen from the evaluation results obtained from this study.
3. Association rules can be a strategy in cross-marketing to get maximum acceptance.

REFERENCES

Annie, L., Mc, C. and Kumar, A. D. (2012) 'Market Basket Analysis for a Supermarket based on Frequent Itemset Mining', *International Journal of Computer Science Issues(IJCSI)*, 9(5), pp. 257–264. Available at: http://www.ijcsi.org/papers/IJCSI-9-5-3-257-264.pdf.

Berndtssom, M. *et al.* (2008) *Guide For Students In Computer Science And Information Systems*. London: Springer.

Berry, M. J. and Linoff, G. S. (2004) *No TitleData Mining Techniques: For Marketing, Sales, and Customer Relationship Management 2nd Edition*. Jhon Wiley and Sons.

Box, G. E. P., Jenkins, G. M. and G.C.Reinsel (1994) *Time Series Analysis: Forecasting and Control*. New Jersey: Upper Saddle River, Prentice Hall.

Chasin, J. R. (2003) 'Implementation of A Cross-Selling Strategy for A Large Midwestern Healthcare Equipment Company', *Department of Psychology, Southern Illinois University at Carbandole*.

Cohen, M. (2004) *Exploiting response models optimizing cross-sell and up-sell opportunities in banking*. Information Systems.

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(1), pp. 118–132.

Gustaf, M. (2009) *Perancangan program Apikasi*

Peramalan Penjualan Berbasis ARIMA dan Neural Network. Jakarta: Program Ganda Teknik Informatika dan Matematika. Universitas Bina Nusantara.

Han, J. and Kamber, M. (2000) *Data Mining: Concepts and Data Mining: Concepts and Techniques: Chapter 6. Mining Association Rules in Large Databases*. Simon Fraser University.

Han, J. and Kamber, M. (2006) *Data Mining Concept and Tehniques*. San Fransisco: Morgan Kauffman.

Kaur, M. and Kang, S. (2016) 'Market Basket Analysis: Identify the Changing Trends of Market Data Using Association Rule Mining', *Procedia Computer Science*. Elsevier Masson SAS, 85(Cms), pp. 78–85. doi: 10.1016/j.procs.2016.05.180.

Kothari, C. R. (2004) *Research Methodology:Methods and Techniques*. New Delhi: New Age International (P) Ltd.

Kusrini (2009) *Algoritma Data Mining*. Jakarta.

Larasati, D. P., Nasrun, M. and Ahmad, U. A. (2015) 'Analisis Dan Implementasi Algoritma Fp-Growth Pada Aplikasi Smart Untuk Menentukan Market Basket Analysis Pada Usaha Retail (Studi Kasus : Pt . X) Analysis and Implementation of Fp-Growth Algorithm in Smart Application To Determine Market Basket Analysi', 2(1), pp. 749–755.

Larose, D. T. (2005) *Discovering Knowledge in Data: An Introduction to Data Mining*. John Willey & Sons, Inc.

Masnur, A. (2015) 'Analisa Data Mining Menggunakan Market Basket Analysis untuk Mengetahui Pola Beli Konsumen', *SATIN-Sains dan Teknologi Informasi*. Available at: http://jurnal.stmik-amikriau.ac.id/index.php/satin/article/download/24/pdf.

Nursanti, I. and Vydia, V. (2013) 'Market Basket Analysis Pada Perusahaan', 11(1), pp. 13–22.

Pressman, R. S. (2010) *Software Engineering : a practitioner's approach*. New York: McGraw-Hill.

Rahmawati, N. *et al.* (2017) 'Aplikasi Data Mining Market Basket Analysis untuk Menemukan Pola Pembelian di Toko Metro Utama Balikpapan Application of Data Mining Market Basket Analysis to DeterminePurchase Pattern at Toko Metro Utama Balikpapan', 8(2012), pp. 1–8.

Venkatachari, K. and Chandrasekaran, I. D. (2016) *Market Basket Analysis Using Fp Growth and Apriori Algorithm: a Case Study of Mumbai Retail Store*, *BVIMSR Journal of Management Research*. Available at: https://www.lib.uts.edu.au/goto?url=http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=115496386&site=ehost-live.

Wu, X. and Kumar, V. (2009) *The Top Ten Algorithm in Data Mining*. New York, USA: CRC Press Taylor & Francis Group, LLC.