MARKET BASKET ANALYST BASED ON WEBSITE USING ECLAT ALGORITHM (CASE STUDY POLA PHARMACY)

Edgar Lawrence [1)], Bagus Mulyawan [2)], Novario Jaya Perdana [3)]

[1)]Jl Jatipadang Street No 9, Pasar Minggu South Jakarta. DKI Jakarta
email: edgarlawrence96@gmail.com[1)]
[2) 3)] Fakultas Teknologi Informasi, Universitas Tarumanagara
Jl. Letjen S. Parman No.1, Jakarta

ABSTRACT

Market basket analyst is a data mining technique to discover associations between datasets. Association rule mining identifies a relationship between a large set of data item. When a large quantity of data is constantly obtained and stored in database, several industries are becoming concerned about mining association rules from their databases. Market basket analysis examines customer buying patterns by identifying associations among various items that customers place in their shopping baskets. It is helpful to examine customer purchasing behaviour and assists in increasing sales. So this system is intended to develop a system for market basket analysis on Pola Pharmacy which generate association rules among itemsets with the use of ECLAT (Equivalence Class Transformation) algorithm. This system supports the decision making process for a market expert.

1. Background

Market basket analysis is a mathematical modelling technique if the customer buys a certain group of items, the customer is likely to buy another group of items, the customer is likely to buy another group of items. It is used to analyse the customer purchasing behaviour, help in increasing the sale transaction data. Market basket analyst technique focused on discovering the purchasing patterns by extracting associations, or co-conccurrences from showroom transactional data. It determines the product which are bought together and to let the employee know which product is potentially will be increase the profit in the future.

Association rule mining is used in customer market analysis. Association rule has recently received much attention from the database community. The process of finding association rules has two separate phases. In the first phase, find all combinations of items that have transaction support above the minimum support count. In the second phase, use the frequent itemsets to gerenate the traditional horizontal database format for mining. In a vertical database, each item is associated with it's corresponding transaction id (TIDset). Mining algortihms using the vertical format have shown to be very effective and usually outperform horizontal approaches because frequent itemsets can be countered via TIDset intersection in the vertical approach.

2. Literature Study

The application of this market basket analyst is given the name "Market Basket Analyst based on Website using Eclat Algorithm (Case Study Pola Pharmacy).

2.1 Association Rules

[1] According to Nurbojatmiko, Eri Rustamaji, Asep Fajar Firmansyah, Association Rules is a method to determine an interesting relationship or pattern between variables in a large set of data. This concept was first introduced in the case of transaction in supermarkets stored on the system namely Point of Sales (POS) to find goods purchased simultaneously by consumers. This method called market basketball analyst. The main purpose of the market basketball analysis is to find the behavior or pattern of shopping from customers in online stores, supermarkets and so on.

The problem that often occurs in the use of this method is that there are often many possibilities that aries from datasets (itemset). The use of this method require a long execution time. In determining an association rule mining there is a measure of attractiveness, such as :

1. Support : this measure use a comparison of a transaction in the dataset (itemset) to the number of all transactions in the dataset. With the equation such as

$$\text{Support} = S(A \cup B)$$

$$\frac{number\ of\ a\ transaction\ with\ contain\ A\ and\ B}{all\ transaction} \quad \text{.....(1)}$$

2. Confidence : this is used to show the strength of the relationship between dataset on an association rules. The equation is as follows :

$$\text{Confidence} = P(B \mid A) = \frac{number\ of\ a\ transaction\ with\ contain\ A\ and\ B}{all\ transaction} \quad \text{.....(2)}$$

2.2 Market Basket Analyst

[2] According to Margareth Rouse, market basket analyst is a data mining technique used by retailers to increase sales by better understanding customer purchasing patterns. It involves analyzing large datasets such as purchase history to reveal product groupings as well as products that are likely to be purchased together. The adoption of market basket analysis was aided by the advent of electronic point-of-sale (POS) systems. Compared to handwritten records kept by store owners, the digital records generated by POS systems made it easier for applications to process and analyze large volumes of purchase data.

Implementation of market basket analysis require a background in statistics and data science, as well as some algorithmic computer program skill.

There are two types of market basket analysis :

1. Predictive market basket analysis : this type considers items purchased in sequence to determine cross-sell
2. Differential market basket analysis : this type considers data across different stores, as well as purchases from different customer groups during different times of the day, month or year. If a rule holds in one dimension (like store) but does not hold in the others, analyst can determine the factor responsible for the exception. These insights can lead to new product offers that drive higher sales.

3. Method

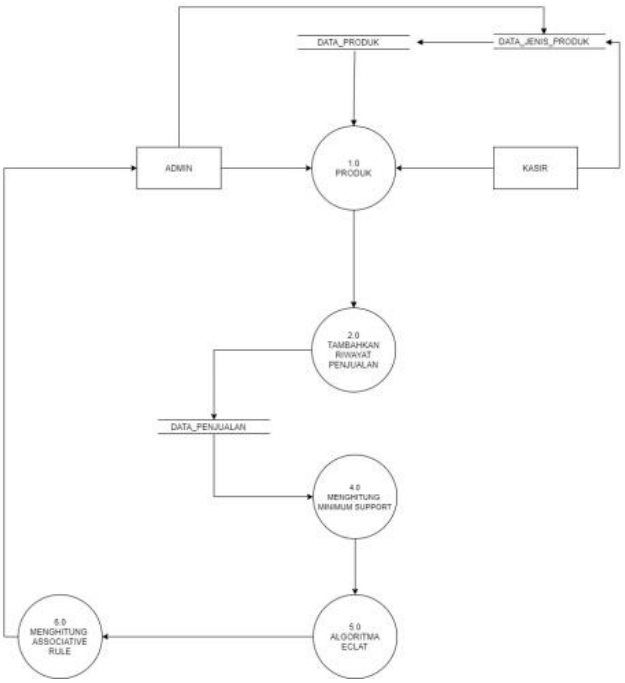
[3] According to Jacek Błażewicz, Wieslaw Kubiak; Tadeusz Morzy, Marek Rusinkiewicz in Handbook on Data Management In Information System, Eclat employs prefix-based classes to reduce the search space. Each itemset are kept sorted in their lexicographic order. The eclat algorithm recursively merges discovered frequent itemsets and uses “tidlists” to evaluate support. When two frequent (k – 1) itmesets are merged to form a candidate k-itemset, their “tidlist” are intersected ro form the “tidlist” of the new candidate.

[4] Eclat algorithm is algorithm that been used for finding frequent itemset into the database. Eclat algorithm searching in depth-first search into database using vertical and if the database is horizontal, then eclat algorithm requiring to converted the database into vertical.

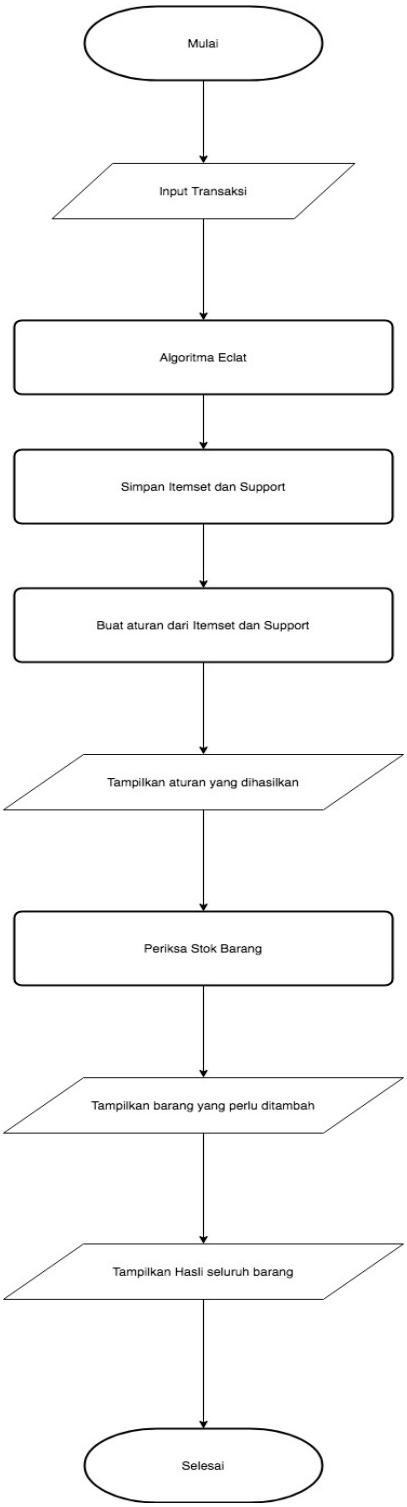
Eclat algorithm had 3 phase :

1. Initiation Phase : Calculation of construction for creating frequent 2 itemset.
2. Transformation Phase : Scheduling the frequent 2 item set to the processor.
3. Asynchronous Phase : Construction of frequent k item set based from user requirement.

In this application, Eclat Algorithm implemented to an analyst and inventory website that been used for analyst every of each customer purchased based on every transaction that can be expalained through DFD and also the flowchart.



Picture 1 DFD (Data Flow Diagram)



Picture 2 Flowchart

4. Result & Discussion

The result from eclat algorithm from testing the data that obtained from Pola Pharmacy based on our agreement from the manager. When testing the data, data use the eclat algorithm based on depth first search method that explained in method section.

Before the data processed into eclat, first the algorithm must convert when the database is horizontal into the vertical. The vertical definition is when the common database usually show how many transaction that appear in in every coloumn on the tabel, eclat will change the structure into every item that contained in every TID list (Table Identifier).

Market basket analyst program works by obtained the whole transactional purchase history in purchased history section that user can input every single of purchased history then the data will be processed and analyze for retrieving the minimum support and association rules.

The conversion will begin like this below of the table

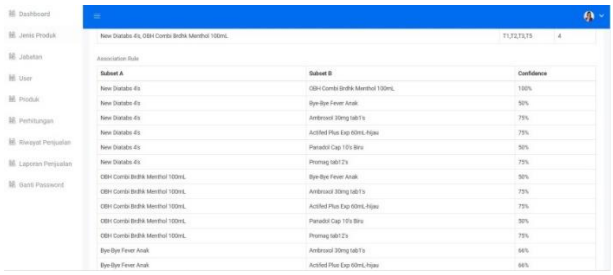
Table 3 Minimum support 3

Item Set	TID	Support
New Diatabs 4's	T13,T12,T16,T14	4
Panadol Cap 10's Biru	T13,T15,T12	3
Promag tab12's	T12,T14,T16	3
OBH Combi Brdhk Menthol 100mL	T13,T14,T16	3
New Diatabs 4's, Promag tab12's	T12,T16,T14	3
New Diatabs 4's, OBH Combi Brdhk Menthol 100mL	T13,T16,T14	3
Panadol Cap 10's Biru	T12,T13,T15	3
Promag tab12's, New Diatabs 4's	T12,T14,T16	3
OBH Combi Brdhk Menthol 100mL, New Diatabs 4's	T13,T14,T16	3
New Diatabs 4's	T12,T13,T14,T16	4

After the minimum support 3 calculation is finished, the algorithm will continiue to calculate the association rule for giving the result to the user of how much percentage of probability of every product that be combined when will be purchased.

Table 2 association rule

Subset A	Subset B	Confidence
New Diatabs 4's	OBH Combi Brdhk Menthol 100mL	75%
New Diatabs 4's	Promag tab12's	75%
New Diatabs 4's	Panadol Cap 10's Biru	50%
OBH Combi Brdhk Menthol 100mL	Promag tab12's	66%
OBH Combi Brdhk Menthol 100mL	Panadol Cap 10's Biru	33%
Promag tab12's	Panadol Cap 10's Biru	33%



Picture 1 Association Rule Result

5. Conclusion

Based on the final results from point number 4 above it can be concluded that eclat algorithm can being used for give product recommendation to user based on purchased history from customer by calculating each product that been purchased in each transaction. Eclat algorithm can be used for helping the user for complete bookkeeping when an user back up transactional history by an digital and online method.

The market basket analyst application will be used as bookkeeping and also analyst of higher profit prediction. The application works depends on the purchased transactional data that every data that been inputted through the system can affect the profit prediction based on the quantity of the data that the user input periodicaly.

REFERENCES

[1] Nurbojatmiko, Asep Fajar Firmansyah, Analysis of Study Program Selection Patterns Using FP-Growth and Eclat Methods, November 2018

[2] Rouse, Margaret. Definition of Market Basket Analyst. https://searchcustomerexperience.techtarget.com/definition/market-basket-analysis, 15 Febuari 2020

[3] Błażewicz, Jacek; Kubiak, Wieslaw; Morzy, Tadeusz; dan Rusinkiewicz, Marek. Handbook on Data Management In Information System. Berlin: Springer-Verlag, 2003

[4] Asror, Ibnu. "Penerapan Metode Association Rule Menggunakan Algoritma Apriori Pada Simulasi Prediksi Hujan Wilayah Kota Bandung". Jurnal Ilmiah Teknologi Informasi Terapan,Volume II, Nomor 2, April 2016.

[5] Sudarsano; Wijaya, Alex dan Andri. "Perbandingan Algoritma Eclat dan FP-Growth Pada Penjualan Barang". Bina Darma Conference on Computer Science, Agustus 2019.

[6] Yanto, Budi; Erni, Rouza dan Jufri. "Strategi Algoritma *Depth-First Search* (DFS) dan Alpa Beta Pruning Pada Permainan COC (*Clash Of Clash*)". Riau Journal Of Computer Science, Vol.3 Nomor.2, Juli 2017.

[7] Kaur, Manjit; Garg, Urvashi; Kaur, Sarbjit. "Advance Eclat Algorithm for Frequent Itemset Generation". International Journal of Applied Engineering Research, November 2015

[8] Ma, Zhiyong; Yang, Juhncheng. "An Improved Eclat Algorithm for Mining Association Rules Based on Increased Search Strategy". International Journal of Database Theory and Application, July 2016.

[9] Yu, Xiaomei; Wang Hong. "Improvement of Eclat Algorithm Based on Support in Frequent Itemset Mining". JOURNAL OF COMPUTERS, VOL. 9, NO. 9, September 2014.

[10] Jia, Linlin; Xiang, Laisheng. "An Improved Eclat Algorithm Based on Tissue-Like P System with Active Membranes". International Journal of Database Theory and Application. September 2017.