Analysis of Shopping Cart in Retail Companies Using Apriori Algorithm Method and Model Profset.

Putri Agung Permatasari[1], Linawati[2] , Lie Jasa[3]

Department of Electrical and Computer Engineering, Post Graduate Program, Udayana University
*agungputri.p@gmail.com@gmail.com

This research needed a priori algorithm method to produce association rules, with a "if-then" pattern. Apriori algorithm uses an iterative approach known as level-wise search, where k-groups of products are used to explore (k + 1) product groups or (k + 1) -itemset (Han and Kamber, 2001). A priori algorithm method is expected to provide vacancy information on one of the stock items that are often purchased simultaneously by the customer. Product optimization is carried out using the profset model, which evaluates the profit margins generated per product and is developed to maximize cross-selling opportunities.
Based on this background, research will be conducted under the title Shopping Cart Analysis in Retail Companies Using the Apriori Algorithm Method and Model Profset.

Index Terms— Algortm Apriori, Model Profset, itemset

I. INTRODUCTION

The information system used by this company has not been able to provide customer shopping habits. The system can not provide information on the relationship between items purchased by customers so that there is often a vacancy of one stock of goods that are often purchased simultaneously. This research the authors raise problems in service that often occurs in retail businesses, namely minimarket owners or retail stores that still tend to place shelf position without regard to the relationship between goods, making it difficult for consumers to find the goods they want (Triyono, 2006). The use of manual methods in analyzing customer shopping habits is difficult, because of the rapid development of data. The application of data mining can help to analyze data obtained from transactions in the information system so that it can explore the patterns of customer shopping habits in the retail company.

This research needed a priori algorithm method to produce association rules, with a "if-then" pattern. Apriori algorithm uses an iterative approach known as level-wise search, where k-groups of products are used to explore (k + 1) product groups or (k + 1) -itemset (Han and Kamber, 2001). A priori algorithm method is expected to provide vacancy information on one of the stock items that are often purchased simultaneously by the customer. Product optimization is carried out using the profset model, which evaluates the profit margins generated per product and is developed to maximize cross-selling opportunities. based on these findings, authors then studied other literature to find other methods that can be used for cost optimization, where research [16] gave an idea to applied data mining

Based on this background, research will be conducted under the title Shopping Cart Analysis in Retail Companies Using the Apriori Algorithm Method and Model Profset.

II. LITERATURE REVIEW

The initial research that supports this research is a study by Handojo, 2007 with the title data mining application to facilitate the association of purchasing goods with the Market Basket Analysis method. And research from Asana, 2013 with the title of shopping basket analysis with a priori algorithm in retail companies. This study provides conclusions about data mining applications that can process transaction data provided by users to find frequent itemets and association conditions that meet the minimum requirements for items in graphics and text. This application discusses the association of goods that are often purchased together with consumers in supermarkets who request this information can give consideration in purchasing decisions to buy goods and arrangements on supermarket shelves.

Research on data mining using the Profset Model has also been conducted by Tom Brijs et al. In this study the framework for optimizing products in supermarket data using the model profset method concluded the need for

additional improvements to the model that will produce and in large shopping basket data. More specifically, it will discuss ways to promote the composition of sets of products with the highest personal profit to increase overall profitability.

The current research will discuss the algorithm used in the previous research. Merging is expected to improve and layout good goods.

A. *Association*

Association The task of association in data mining is to find attributes that appear at a time. In the business world more commonly called shopping cart analysis. Examples of associations in business and research are: a. Examining the number of subscribers from mobile telecommunication companies that are expected to provide a positive response to the offer of service upgrades provided. B. Determine items in supermarkets purchased simultaneously and which are never purchased simultaneously. [1]

Analysis methods in data mining can be classified by: ;[3]

1.Association Search association rules that show the conditions of attribute values that often occur together in a set of data. Associated analyzes are often used to analyze basketball markets and transaction data.

2.Classification and prediction The process of finding models (functions) that explain and distinguish classes or concepts, with the aim that the model obtained can be used to predict which class or object has an unknown class label.

3.Clustering

Analyze data objects where unknown class labels can be used to determine unknown class labels by way of grouping data to form a new class. Maximizing intra -class similarity and minimizing the interclass resemblance

B. *Data Mining*

Data mining is the process of data analysis using software to find patterns and rules (rules) in the set of data. Data mining can analyze large amounts of data to find the knowledge to support decision making [4]. Along with the increasing computerization of community activities, greater data storage capacity and faster computer processing capabilities,[2] resulting in data generated and stored by individuals and organizations every day is very large.[8] Data mining are activities extract or mine knowledge from large amounts of data, this information will be very useful for the development [6]. Data Mining is the term used to describe the process extract value / information from the database.[17] .Data Mining is known as Knowledge Discovery in Database (KDD) [12]. There are four things are needed in order to effectively data mining such as :high quality data, right of data, examples of which are adequate, and the correct device [10].

III . METHODOLOGY

The flow of research is as follows defining the problem of the system created. Data collection and literature studies related to shopping basket analysis using a priori algorithms. As well as studying and understanding the processes that occur in the system to be made so that system modeling can be done. Designing a data mart for the shopping cart analysis process. Application development includes software design that is used as a user interface shopping cart analysis using a priori model algorithm algorithm in the Visual Basic .Net 2008 platform. System testing and system testing analysis. The last process is to do the shopping basket analysis process, so that it can be seen the relationship between goods that are often purchased by customers and make conclusions.

The following is attached Chart 3.1 Research Flow.

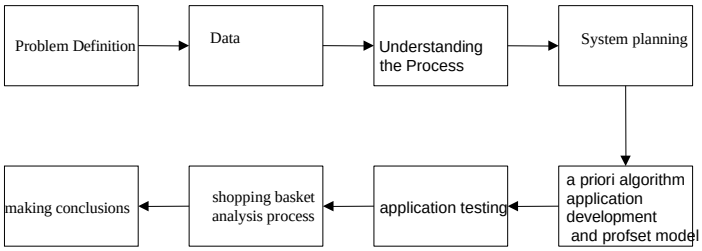

Diagram 3.1 Schedule of Activities

C. *Algortm Apriori*

Apriori algorithm includes the type of association rules on data mining. In addition to Apriori, which is included in this algorithm is the method of Generalized Rule Induction and Hash Based algorithm. The rules that state the association between several attributes are often called affinity analisys or market basket analysis [1].

For the application of a priori algorithms, data structures are generally needed to store frequent itemset candidates for an iteration to k and to store the resulting frequent itemset.[4].

D. *Clustering*

Clustering can be said as identification of smilar classes of objects [11. The clustering is different from the classification of the absence of the target variable in the clustering. Clustering does not attempt to classify, estimate, or predict the value of the target variable. However, the clustering algorithm tries to divide the entire data into groups that have a similarity record in one group will be maximal, while thsimilarity with records in other groups will be of minimal value [14].

Table 3.1 Sales Transaction Data

Transaction Code	Food Material
10041	Milk, Tea, Sugar
10042	Tea, Sugar, Bread
10043	Tea, Sugar
10044	Milk, Bread
10045	Milk, Sugar, Bread
10046	Tea, Sugar
10047	Sugar, Coffee, Milk
10048	Sugar, Coffee, Milk
10049	Milk, Bread, Coffee
10050	Sugar, Tea, Coffee

Table 3.1 is a sample of sales transaction data for a retail company. In the table there is a combination of items purchased by the customer. Table 3.1 will be used as a reference to find a high frequency pattern (frequent pattern mining), namely the discovery of associations and correlations between items that are in the amount of transaction data. The minimum support value used is 2. The first step is to scan sales transaction data to find a 1-itemset candidate with the support value of each itemset.

Table 3.2 1-itemset candidates

Itemset	Support
Milk	6
Tea	5
Sugar	7
Bread	4
Coffee	4

The next step is to find a 1-itemset, that is, an itemset that has a support value greater than or equal [9] to 2 in table 3.2.

Table 3.3 1-*itemset*

Itemset	Support
Milk	6
Tea	5
Sugar	7
Bread	4
Coffee	4

1-itemset has been found (Table 3.2), then found 2-itemset by searching for 2-itemset candidates by scanning transaction data.

Table 3.4 2-itemset candidates

Itemset	Support
Milk, Tea	1
Milk, Sugar	4
Milk, Bread	3
Milk, Coffee	3
Tea, Sugar	5
Tea, Sugar	1
Tea, Coffee	1
Sugar, Bread	2
Sugar, Coffee	3
Bread, Coffee	1

Then the 2-itemset candidate is scanned to find 2-itemset that meets the minimum support requirements.

Table 3.5 2-*itemset*

Itemset	Support
Milk, Sugar	4
Milk, Bread	3
Milk, Coffee	3
Tea, Sugar	5
Sugar, Bread	2
Sugar, Coffee	3

Next search for 3-itemset candidates based on 2-itemset data. In a priori algorithm [7], each member of the itemset must also be an itemset that meets the minimum support. So the candidates that have been found as in table 3.5 must be tested by members of each itemset candidate, this process is known as prune.

Table 3.6 3-itemset candidates

Itemset	Support
Milk ,Sugar, Bread	1
Milk, Sugar, Coffee	2
Milk, Sugar, Tea	1
Tea, Sugar, Milk	1
Tea, Sugar, Coffee	1

Table 3.6 is a candidate for pruning results.

Table 3.7 3-itemset candidates After Pruning

Itemset	Support
Milk, Sugar, Bread	1
Milk Sugar, Coffee	2

Then look for the itemset that meets the specified minimum support.

Table 3.8 3-*itemset*

Itemset	Support
Milk, Sugar, Coffee	2

E. ***Profset Model***

The profile model was developed to maximize cross-selling opportunities by evaluating the profit margins produced per set of products. This is caused by fatherly evaluating the business value is less than optimal if only taking into account the profits generated by each product. To optimize business profits, it is important to consider the set of products that are often purchased.Systems Development Engineering [5].

The PROFSET model was developed from the product selection model, namely:[5]

Keterangan:

T_j = Individual transaction for-j

SP_i = selling price of the product i

PP_i

f_i = product purchase price i

=number of products i bought during the transaction T_j

m_{Tj} = gross profit derived from the transaction T_j

M_X = gross profit from the product set X

The updated and improved PROFSET model is as follows:

Objective function [5]

Max

Limitation:

$$\sum_{c=1}^{n} \sum_{i \in C_c} Q_i = ItemMax \quad (1)$$

$$\forall X \in L, \forall i \in X: Q_i \geq P_x \quad (2)$$

$$\forall C_c: \sum_{i \in C_c} Q_i \geq ItemMin_{C_c} \quad (3)$$

Information:

L = collection of product sets

Kategori $C_1, \ldots, C_n$ = set

$P_x, Q_i \in \{0, 1\}$ is a decision variable that will be optimized

$Cost_i$= total cost (storage + handling) for the product i

Calculation:

Taken from Table 3.5 2-itemset and taken the greatest Support

Table 3.9 *2-itemset*

Itemset	Support
Milk, Sugar	4
Milk, Bread	3
Milk, Coffee	3
Tea, Sugar	5
Sugar, Bread	2
Sugar, Coffee	3

The data that the basic combination will look for is as follows:

X_1 = { Milk, Sugar }

X_2 = { Sugar, Coffee }

X_3 = { Sugar, Tea }

Table 3.10 Price

Itemset	Purchase price	Selling price	Storge and Handling fees
Milk	$25	$35	$3
Sugar	$14	$18	$2
Coffee	$10	$14	$1
Tea	$3	$5	$1

Information :

Handlling = Moving goods from the producer or distribution to the warehouse

Storge = price of maintenance of goods during storage

Finding Mx from the above combination is taken from the following formula:

1. M(X_1) = ($35 – $25) + ($18 – $14) = $14
2. M(X_2) = ($18 – $14) + ($14 – $10) = $8
3. M(X_3) = ($18 - $14) + ($5 - $3) = $6

A possible combination is 3C4:

Known:

A = Milk

B = Sugar

C = Coffee

D = Tea

1. Z_1 = A, B, C
2. Z_2 = A, B, D
3. Z_3 = A, C, D
4. Z_4 = B, C, D

Then :

1. Z_1 = Milk, Sugar, Coffee
2. Z_2 = Milk, Sugar, Tea
3. Z_3 = Milk, Coffee, Tea
4. Z_4 = Sugar, Coffee, Tea

Put in the formula:

$$Z = (M(X_1)PX_1 + M(X_2)PX_2 + M(X_3)PX_3 - C_1Q_1 - C_2Q_2 - C_3Q_3 - C_4Q_4)$$

1. Z_1 = ($14(1) + $8(1) + $6(0) – $3(1)- $2(1)-$1(1)-$5(0) = $16
2. Z_2 = ($14(1) + $8(0) + $6(1) – $3(1)-$2(1)- $1(0)-$1(1) = $14.5
3. Z_3 = ($14(0) + $8(0) + $6(0) – $3(1)-$2(0)- $1(1)-$1(0) = -$4
4. Z_4 = ($14(0) + $8(1) + $6(1) – $3(0)-$2(1)- $1(1)-$1(1) = $10.5

From the data above for more optimal goods procurement is

1. Z_1 = Milk, Sugar, Coffee

With the total profit to be gained when the combination is $16

F. ***Use Case Diagram***

Use case diagrams that illustrate how people interact with the system. A use case represents an interaction between actors and the system. Use case is a technique for recording the functional requirements of a system. The use case describes the typical interaction between the users of the system and the system itself, by, giving a narrative about how the system is used (Fowler, 2005) [28]. The following is the use case diagram of the shopping basket analysis application in table 3.11, which has a basic case including Log in, managing a login account, shopping basket analysis, and importing data. And the incluent case includes showing the association rule and PROFSET model analysis.continues with planning, modeling, construction, and submission. The stage of the waterfall approach model can be seen in Figure 3. 11

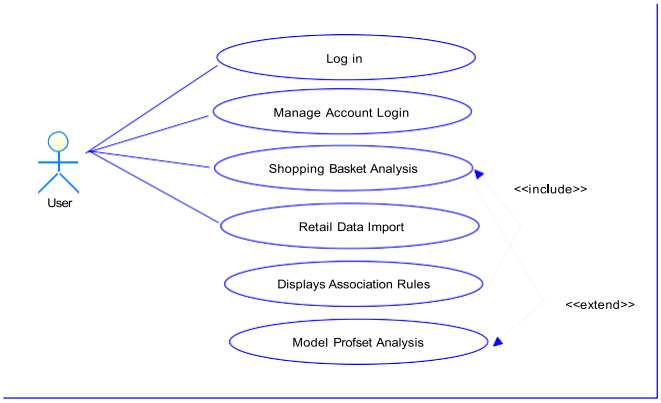

Figure 3.11 Use Case Cart Analysis Diagram

Cart Analysis Diagram, the user has entered the application and the system has displayed the Main Form. First the user presses the Shopping Cart Analysis button. The system then displays the Initial Analysis Form and the user then presses the Next button to start the process. After that the system then displays the Support & Confidence Form. In the Support & Confidence Form the user enters a minimum value of support and minimum confidence then presses the Next button.

III. RESULT

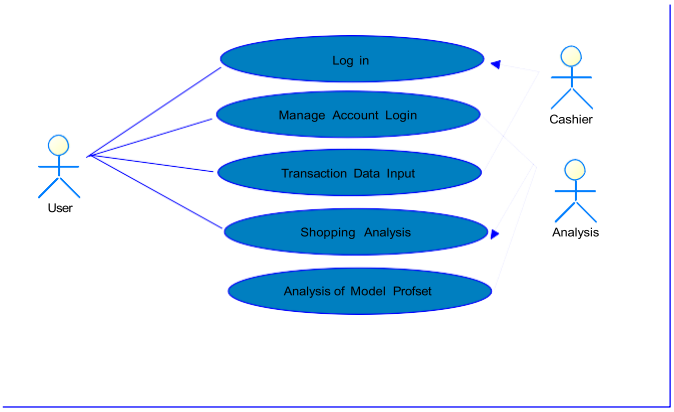

Figure 3.12 Use Case Shopping Cart Analysis Application

D. ***Activity Diagram***

Diagram, the user has entered the application and the system already displays the Main Form. First the user presses the Shopping Cart Analysis button. The system then displays the Initial Analysis Form and the user then presses the Import Data transaction button to start the process. After that the system then displays the desired Transaction Data, then press the further button then displays the Support & Confidence Form. In the Support & Confidence Form, the user enters the minimum support and minimum confidence values and then presses the Next button. After entering that value, the system then displays.

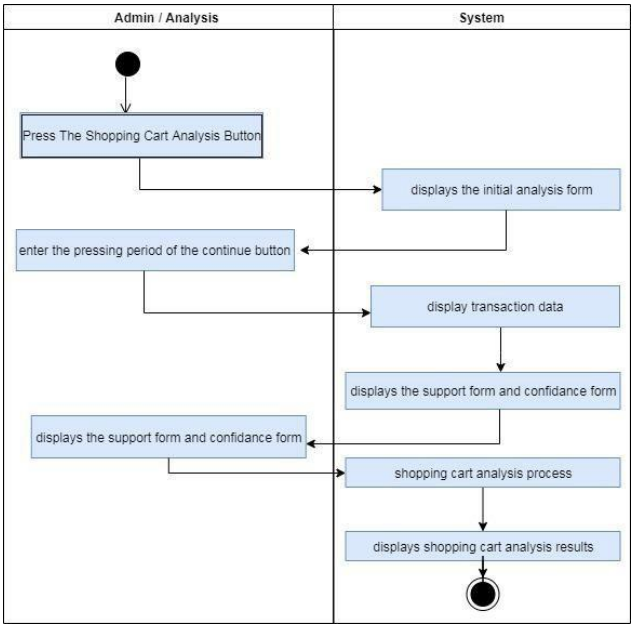

Figure 3.13 Activity Diagram of a shopping basket analysis

In figure 3.14 Activity Diagram Analysis Model PROFSETstarting from the system of Shopping Cart Analysis results, from the system. The user presses the PROFSET Model button after which the system carries out the PROFSET Model Analysis process and then the system displays the PROFSET Model Analysis Results.

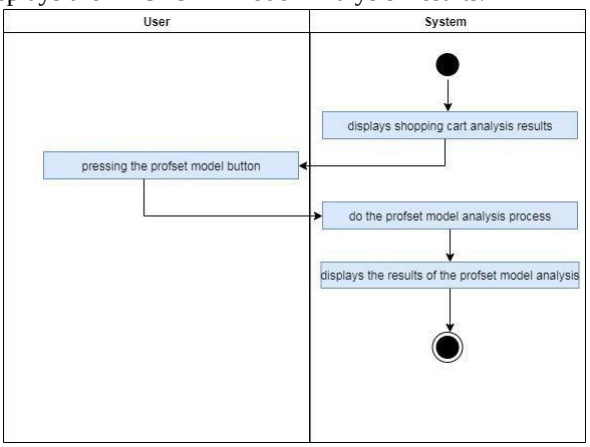

Figure 3.14 Activity Diagram Profset Model

E. ***Sequence Add Data Diagram***

In Figure 3.15 Sequence Diagram of Shopping Cart Analysis illustrates the sequence of events in use case Shopping Cart Analysis with scenarios begins with the user starting the analysis on frmAnalisis. getItemset message to the sales_det table object. From the table of sales_det gives a reply message that is the data requested in accordance with the input to ctrlAnalisis. Then from ctrlAnalisis. After the data is correct it will display the Shopping Cart Analysis
Results in the shopping cart analysis results

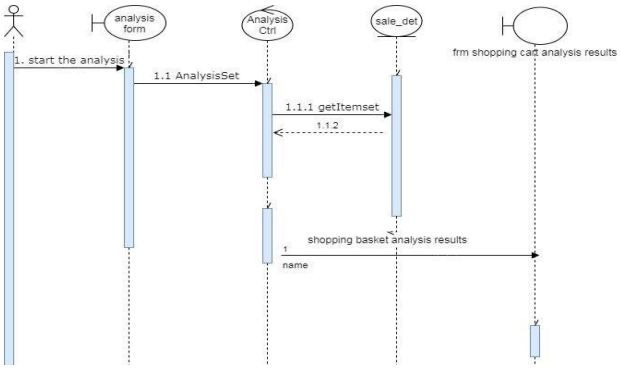

Figure 3.15 Sequence Chart of Shopping Cart Analysis

In Figure 3.16 Sequence Diagram Analysis of the PROFSET Model From frmAnalysis Shopping Cart The user carries out the process of analyzing the profset model.

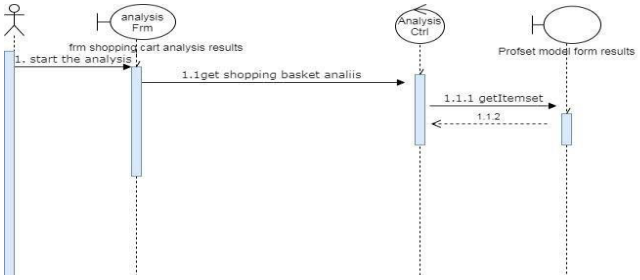

Figure 3.16 Sequence Diagram Analysis of the PROFSET Model

F. System Interface Design

The following is a system interface design that will be used in developing market shopping basket analysis applications. Log In Form Design At the start of the application, the system will display FormLog In (Figure 3.17). In this form the user enters the user id and password.

from this existing relationship, it will be able to analyze

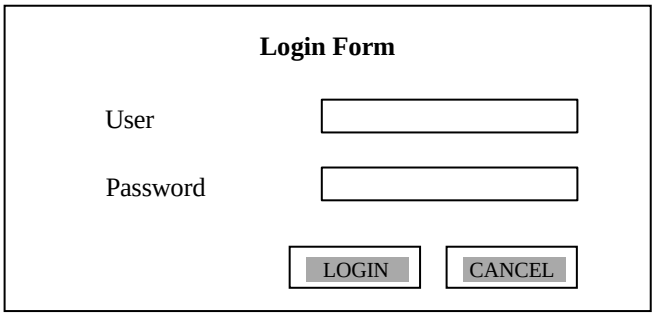

Figure 3.17 Login Form

The login form will appear when the system is first run. The login form functions to authenticate the username and password. In the login form there are two text boxes namely a text box for entering a user name and a text box for entering a password. In addition there is also a button to login. The login form display is shown in Figure 3.17

The main menu will be seen in Figure 3.18 With the rights to see this menu are the admin and analysis.

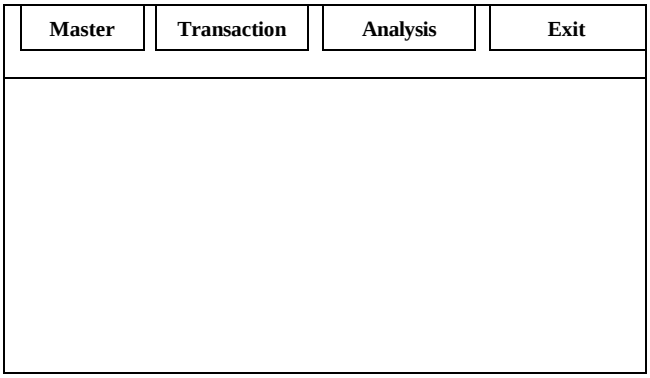

Figure 3.18 Main Menu Form

The Shopping Cart Analysis menu form will appear after the user successfully imports the Transaction Data. In this analysis form the user will input Minimum Support and Minimum Confidence%. After inputting the user does the Shopping Cart Analysis process, this form is also provided back or return to the next form or exit the application. The contents of this form are combinations, association rules, support, and confidence. The appearance of the Shopping Cart Analysis Data Import form is shown in Figure 3.18 below.System Overviev

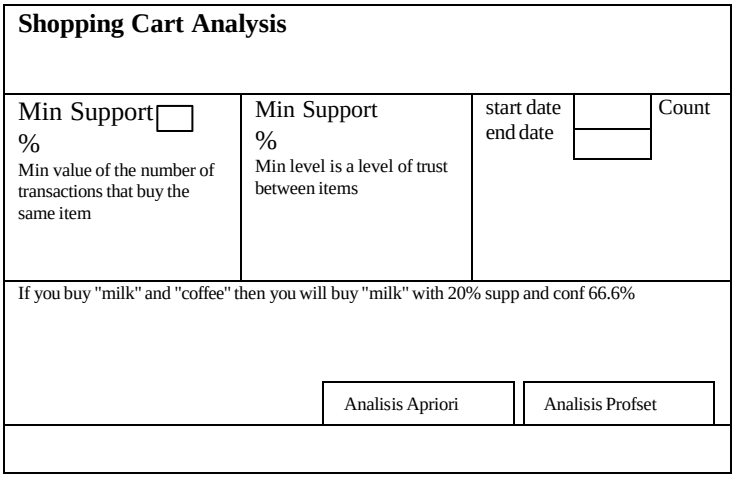

Figure 3.19 Shopping Cart Analysis Form

The Form Analysis Model PROFSSET menu will appear after the user successfully analyzes the Shopping Cart. In this Analysis form the user will not be able to input Minimum Support and Minimum Confidence%. The application will display the existing combination and the price Profist on the item. PROFSET Model Analysis Data is shown in Figure 3.19 below.

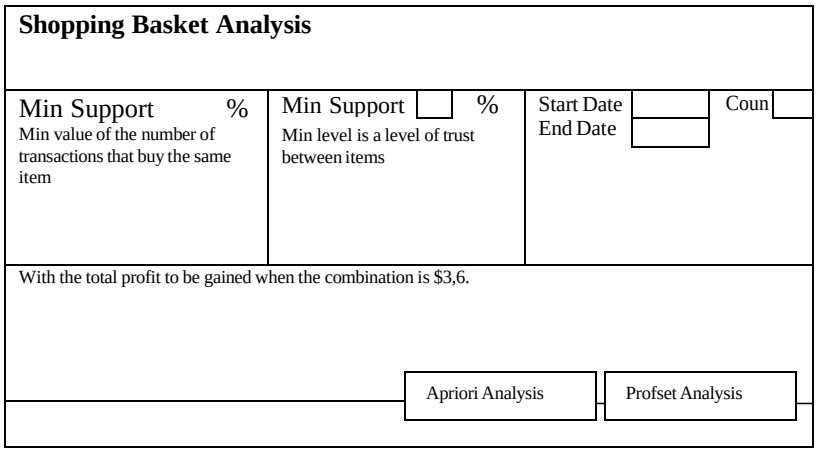

Figure 3.20 PROFSET Model Analysis Form

H. Implementation

The login page is the page that first appears when the system starts. On the login page there is a column for entering the login username and password so that the user can enter the system to access the menu on the main menu, to log in, distinguished access rights, namely, admin, cashier, and analysis. The display of the login page can be seen in Figure 4.1.

4.1 Login Page

The main menu page is the first page that appears when the login is successful, the main menu display below is the rights of the admin and Analysis. On this main menu page there is a master (there is an addition of employee data or manage account logins, transaction data, Analysis and Exit. Gambat can be seen in Figure 4.2.

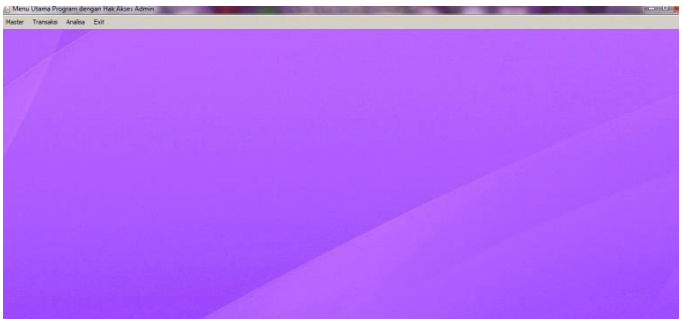

Figure 4.2 Main Menu Page

The master page and employee menu are pages when the admin wants to change password and add user names in accordance with the data access rights. On the employee page there are new (addition of user data), update (update new data), delete, and exit Figure 4.3

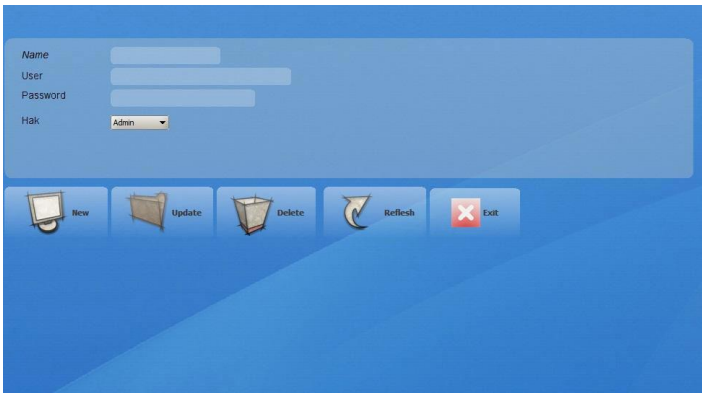

Figure 4.3 Employee Form

The Analysis Form page is the page when the admin or analysis wants to see the results of the analysis when the user chooses the start and end date to determine the transaction data. After that click analysis will display the analysis results. After selecting the period, then enter the min support and min confidence figures in Figure 4.3 and 4.4

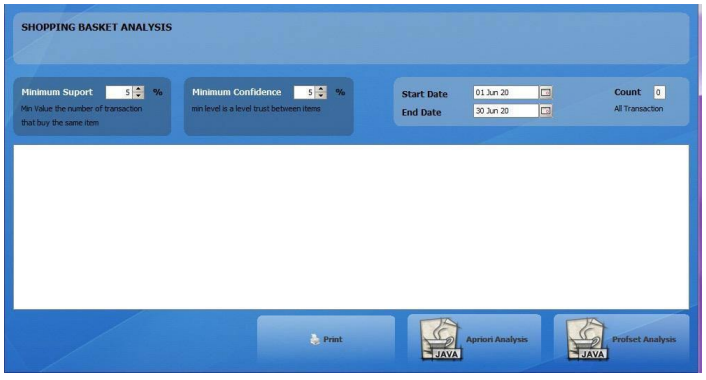

Figure 4.4 Apriori Algorithm Analysis Form

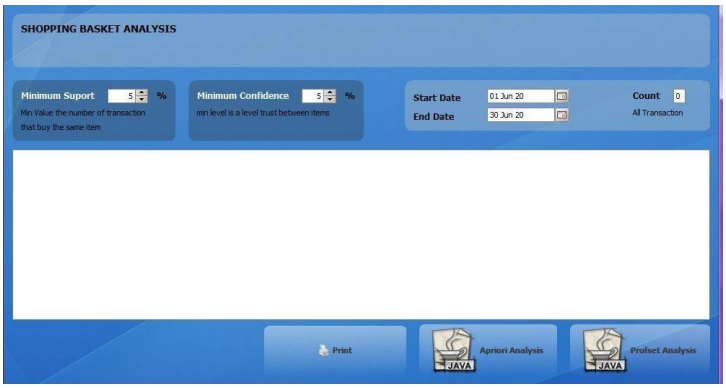

Figure 4.5 Form Analysis Analysis Model Profset

IV CONCLUSION

The process of determining the purchase pattern can be done using data mining with the Appriori algorithm method and seeing the profit margins with the Profset Model. With this method determining the purchase pattern can be done by looking at the choice of consumers to buy goods based on a combination of items. The application of Apriori Algorithms in the Data Mining technique is very efficient and can accelerate the process of forming the combination of sales items. Can help valuable information in making decisions to prepare stock of goods needed by the market. Thereby reducing the expiration rate on goods and data that have been put to good use. and can see a more favorable combination.

This is the hypothesis of further research to apply other than Apriori algorithms can be used to improve the algorithm used can be done to compare Apriori algorithms with Other Algorithms

REFERENCES

[1] Adi Panca, Kheri Arionadi, and Saputra, Komang Oka .2017. "Analysis Shoping Cart at Drugs Store By Using an Apriori Algorithm". *International Journal of Engineering and Emerging Tercnology*, Vol 2,No1, p-issn: 2579-5988.

[2] Ansari Muhammad, Suryadharma I Putu, and Sastra Nyoman Putra.2017."Application of Costomer Clustering Mining in Household with Fuzzy Multi Criteria Decision Making (FMCDM). *Journal of Engineering and Emerging Tercnology*, Vol. 2, No 1, p-ISSN:1693-2951.

[3] Ardiada Dwi, Ariawan Pasek Agus, and Sudarma Made. 2018. " Evalution of Supporting Work Quality Using K-Means Algorihm". *Journal of Engineering and Emerging Tercnology*,Vol.3 , No1, p-ISSN:1693-2951.

[4] Arimbawa Dewa Agung, Prawaba I Nyoman Angga, and Mertasana Putu Arya. "Implementation of Apriori Algorithm Determining Tourism Visit Pattern to Bali ". *International Journal of Engineering and Emerging Tercnology*, Vol. 4, No , p-ISSN:2579-5988.

[5] Brij,Tom dkk. Kerangka untuk Mengoptimalisasi Produk dalam Supermarket dengan Metode Model Profset.University Center, Limburg

[6] Dewi Indah Cahya, Gautama Bara Yuda, and Mertasana Putu Arya.2017. "Analysis of Clustring for Grouping of Productive Industry by K-Medoid Method" *Journal of Engineering and Emerging Tercnology*. Vol. 2, No1, p-ISSN:2579-5988.

[7] Gama, Darma Putra, dan Agung Bayupati. 2016. Implementasi Algoritma Apriori Untuk Menemukan Frequent Itemset dalam Keranjang Belanja. *Journal of Engineering and Emerging Tercnology*, vol 15,No.2, p-ISSN:1693-2951.2016

[8] Hendrayana, Rifky Lana , and Sudarma Made.2017. Application of Neural Network Overview in Data Mining. *Journal of Engineering and Emerging Tercnology*, Vol 2, No. 1, p-ISSN:2579-5988.

[9] IGA Indrawan, M Sudarma, And L Jasa.2018. Implementasi Algoritma FP-Growth dengan Closure Table untuk Penemuan Frequent Itemset pada Keranjang Belanja. *Journal of Engineering and Emerging Tercnology*, vol. 17, No.2, p-ISSN:1693-2951.

[10] Mamang Weking Philipus N, Parta IGN, Weking Antonius Ibi. 2019. "Application of Data Mining With Support Vector Machine (SVM) in Selling Prediction Trend of Spiritual Goods (Case Study : PT. X Bali)." *International Journal of Engineering and Emerging Tercnology*,Vol. 4, No1, p-ISSN 2579-5988.

[11] Mahendra IGN, Suta IB Leo, and Sudarma M.2019. "Classification of Data Mining With Adaboost Method in Determining Credit Providing For Costomers". *Journal of Engineering and Emerging Tercnology*, Vol. 4, No 1, p-ISSN: 2579-598.

[12] Maricar M. Azman, Widiadnyana Putu, and Wijaya Wayan. 2017. "Analysis of Data Mining for Forcesting Total Goods Delivery With Moving Average Methode". *Journal of Engineering and Emerging Tercnology*, Vol. 2. No ,1, p-ISSN: 2579-5988.

[13] Peling IB, Arnawan Inyoman, and Janardana IGN. 2017. " Implementation of Data Mining To Predict Period of Students Study Using Naïve Bayes Algorthm". *Journal of Engineering and Emerging Tercnology*,Vol. 2, No 1, p-ISSN:1693-2951.

[14] Putra Putu Bagus, Suryani Ni Putu, and Aryani Sri. 2018. "Analysis of Apriori Algorthm on Sales Tr ansaction to Arrange Placement of Goods on Minimarket". *Journal of Engineering and Emerging Tercnology*, vol 3, No.1, p-ISSN: 2579-5988.

[15] Putu Angelina, and Made Sudarma. 2018. "Implementation of EM Algorthm in Data Mining for Clustering Famele Coorporative". *Journal of Engineering and Emerging Tercnology*, Vol.3, No1, p-ISSN: 2579-5988.

[16] Pramana I Wayan Surya, Iswardani Putu Risanti, an Aryani. 2019." Application of Data Mining in Optimazation of Hotel „s Foods and Beverage Costs" *Journal of Engineering and Emerging Tercnology.* Vol, 4, No 1, p-ISSN:2579-5988.

[17] Surya Hendrayana PB, Rahardian, and Sudarma Made.2017."Application of Neural Network Overview in Data Mining". *Journal of Engineering and Emerging Tercnology*, Vol. 2, No 1, Vol.3 , No1, p-ISSN:1693-2951.