BASKET MARKET ANALYSIS USING R-BASED APRIORI ALGORITHM TO FIND INFORMATION FROM SALES DATA

Indra Yustiana [a,1,*], Irvan Sutarha [a,2], Nindia Maulawati [a,3], Ilham Maulana Yusuf [a,4]

[a] Departement of Informatic Engineering, Nusa Putra University, Jl. Raya Cibolang Kaler No.21, Kab. Sukabumi 43152, Indonesia
[1] indra.yustiana@nusaputra.ac.id*; [2] irvan.sutarha_ti17@nusaputra.ac.id; [3] nindia.maulawati_ti17@nusaputra.a c.id;
[4] Ilham.maulana_ti17@nusaputra.a c.id
* Corresponding Author

ABSTRACT

Market Basket Analysis is a data mining technique that is used to determine which products a customer will buy simultaneously by analyzing a list of customer transactions. By knowing these products, an e-commerce system can create or develop a customer profile system and can determine its own customer catalog layout. This journal discusses data mining techniques, with association rules that can help check customer buying behavior and increase sales. The result can provide reference prices for cross selling, designing promotions and placing merchandise in stores increasing sales
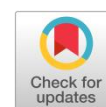
KEYWORDS
Market Basket Analysis
Apriori
Clothing Retail Stores

1. Introduction

Many methods are used in data mining, namely estimation, prediction, classification, clustering, association.[1]

The Apriori algorithm is known as the most established and simplest algorithm but requires a long computation time and requires a large memory allocation to search for itemsets. This is due to repeated scanning of the data. This algorithm is used to determine the correlation between the goods that are of interest to consumers stored in the database. After obtaining frequent itemsets, a rule is taken and then analyzes the comparison of execution speed, rule formation process and rule accuracy from the algorithm.[2]

Based on the problems and considerations above, a research can be carried out on the performance of the Apriori algorithm to determine market basket analysis on beauty products and adult clothing with the title "Implementation of Market Basket Analysis using the Apriori Algorithm based on R Language at Clothes Stores".

The research was made by M Jundi Hakim and Yuma Akbar with the research title "MARKET BASKET ANALYSIS USING APRIORI ALGORITHM BASED ON BAHASA R (Transretail Indonesia Case Study)". Market Basket Analysis are processing technique for discovery of relationships between various kinds of goods. The main Objective of the Market Basket Analysis in retail to provide information for distributors to find out the purchasing behaviour of customers, which can assist distributors in creating the right choice. There are various the algorithm is available to conduct market basket analysis. This journal discusses data mining techniques, with association rules that can help examine customer buying behaviour and held increase sales. The results can provide price references for cross selling, designing promotions and a placing merchandise in stores to increase sales.[3]

ijeat@nusaputra.ac.id

Further research from Imroatun Qoni'ah and Adhie Thyo Priandika with the research title "BASKETBALL MARKET ANALYSIS TO DETERMINE THE ASSOCIATION OF RULES WITH A PRIORI ALGORITHMS (CASE STUDY: TB. TOWER)". Market Basket Analysis is defined as an itemset that is purchased simultaneously by a customer in a transaction. Besides that, it is also used to analyze the pattern of consumer spending. By utilizing processed sales transaction data to obtain information from the transaction data, TB. The tower is a business that is engaged in the sale of building materials and trading tools, which is located at Punggur, Central Lampung. This store does not yet know the patterns of consumer shopping in the shopping cart. The algorithm used is the a priori algorithm because this algorithm reduces the number of candidate itemsets at the start. From the results of this study, it was found that the best-selling item for 1-itemset was 48% Holcim cement. Items for 2-itemset, namely ceramics and holcim cement by 19%. The association rule is that when consumers buy Asbestos Rubber they will buy Asbestos with a confidence value of 94%, When consumers buy Asbestos they will buy Paku Payu with a confidence value of 88%, When consumers buy ceramic trim, they will buy ceramics with a confidence value of 89%, When consumers buy asbestos, you will buy asbestos rubber with a confidence value of 92%. [4]

From the research made by Thomas Brian and Ardhi Sanwidi with the research title "IMPLEMENTATION OF A PRIORI ALGORITHMS FOR MARKET BASKET ANALYSIS BASED R". More sales transactions hence required a system to generate important information. This innovation will solve many problems in the field of sales marketing and inventory, because the products that are not so salable if paired properly will increase the value of sales. However, looking for associations requires a complicated process because of the problem of large product combinations let alone if the retail business has thousands of products. Apriori is a data mining algorithm to find relationships between items on market basket analysis. By finding the pattern of sales transactions is expected to increase business value. In the process undertaken in this research is implemented using R with a priori function to process data. It starts from reading the dataset until it finds a recommendation from the system that has been created using the function in R. Determine the support, confidence and lift values to find the best itemset for the next sale. Trial that has been done with transaction dataset shows best result on value filter support = 0,1 confidence = 0,8 and lift > 1.[5]

- Knowledge Discovery in Database (KDD) and Data Mining

KDD is not a nontrivial process in extracting implicit data that has not been known before, and has the potential to be useful information (Fayyad, 1996). Nontrivial because some of the searches or inferences involved are not a direct computation of a predefined quantity, such as computation of the mean value of a set of numbers. Finally, these patterns must also be understood and understood, although there is the possibility that they cannot be directly and must go through several processes first

information mining could be a arrange within the KDD handle which comprises of the application of information examination and revelation calculations, which can be acknowledged inside the limits of computational productivity, coming about in an count of certain designs (or models) of information. The KDD handle includes employing a database amid the determination, beginning handling, subsampling, and change as required.[6]

Application of Association Methods Using A priori Algorithms In Consumer Shopping Pattern Applications This ponder was conducted in arrange to form a Advertise Wicker container Examination by utilizing Association Rules. The information utilized within the ponder are the deals information of any general store gotten from the Vancouver Island University website. Information were analyzed within the Weka program employing a information set containing 225 distinctive items. Apriori and FP Development, which are Affiliation Rules calculations, were attempted in arrange. Since the information set is categorical, the Apriori calculation did not surrender any comes about. Hence, the FP Growth calculation was utilized and the best 10 rules were given concurring to the conviction esteem. The finest run the show in like manner; a client who buys Drain, Sweet Savor and Pepperoni Pizza (Frozen) also gets eggs. Best run the show with 21.06 Conviction and 1 (100%) certainty values are this run the show. 24 clients who received these 3 items within the dataset gotten eggs. Additionally, too   other

rules were translated in this consider. As a result, item situation within the grocery store can be made agreeing to these rules. In this way,.[11]

Market basket analysis with association rules, **Communications in Statistics** ,Data mining might be a orchestrate inside the KDD handle which comprises of the application of data examination and disclosure calculations, which can be recognized interior the limits of computational efficiency, coming almost in an check of certain plans (or models) of data. The KDD handle incorporates utilizing a database in the midst of the assurance, starting taking care of, subsampling, and alter as requiredConsumer buying design investigation utilizing apriori affiliation , Objective:Cross selling examination is an imperative expository instrument within the retail industry.The deals of a retail store can be made strides by deciding the situating of merchandise, and planning deals advancement plans based on item movement.The five month's charging information sets of a Retail Super Advertise based at Trichy, Tamil Nadu, were considered for the study.Methodology: The issues of item situating in a well-established retail store are inspected utilizing information mining to recognize the thing sets that are bought frequen.[12]

Association Rule Mining is a data mining designed to discover the real connections of data items in transaction data build on associativity. The technique utilizes the Apriori Algorithm to discover association rules. Furthermore, the Apriori Algorithm widely treated to discover frequent itemsets in transaction data. This study aims to enhance the efficiency of the Apriori Algorithm in the mining of association rule as a reference to identifying mixed item deals as a regular promo to offer to customers build on item frequency of buying. The study shows that Association rule mining implementation through enhanced Apriori Algorithm generates results at a higher performance rate or lesser runtime rate compared with the original Apriori Algorithm, and it helps the organization in selecting customer product deals. Anticipated in business flow, the study produced a list of package items for consumers based on strong rules generated by association rule mining at a lesser runtime rate[13]

A new optimization model for market basket analysis with allocation considerations: A genetic algorithm solution approach These days advertise wicker container examination is one of the interested investigate ranges of the information mining that has gotten more consideration by analysts. But, most of the related investigate centered on the conventional and heuristic calculations with restricted variables that are not the as it were persuasive variables of the wicker container advertise examination. In this paper to productive modeling and investigation of the showcase wicker container information, the optimization demonstrate is proposed with considering allotment parameter as one of the imperative and useful variables of the offering rate. The hereditary calculation approach is connected to unravel the defined non-linear twofold programming issue and a numerical illustration is utilized to demonstrate the displayed show. The given comes about uncover that the gotten arrangements appear to be more practical and appropriate. [14]

A Study on Market Basket Analysis and Association Mining In this think about, we connected the showcase bushel investigation to plan office format of an entertainment arcade in Surabaya. The issue faces by the beguilement arcade is clients as it were play in certain recreations that causes numerous amusement machines to be sit out of gear. This issue will be troublesome to resolve since of income design has not been recognized. Subsequently, showcase wicker container investigation is connected to know the client behavior in playing the diversions. As the result we proposed two formats. To begin with proposition format will be planned based on amusement sort. This format will classify amusement machines based on advertise bushel examination comes about in each category where each category is free of other category. The free presumption within the first layout is discharged within the moment format proposition. Within the moment format proposition each amusement category is subordinate of other category. As the result, the moment proposition is more likely to be connected, since this course of action does not taken a toll any cash and does not require particular mater.[15]

Designing facility layout of an amusement arcade using market basket analysis In this study, we applied the market basket analysis to design facility layout of an amusement arcade in Surabaya. The problem faces by the amusement arcade is customers only play in certain games that causes many game machines to be idle. This problem will be difficult to resolve because of revenue pattern has not been acknowledged. Therefore, market basket analysis is applied to know the customer behavior in playing the games. As the result we proposed two layouts. First proposal layout will be designed based on game type. This layout will classify game machines based on market basket analysis results in each category

where each category is independent of other category. The independent assumption in the first layout is released in the second layout proposal. In the second layout proposal each game category is dependent of other category. As the result, the second proposal is more likely to be applied, since this arrangement does not cost any money and does not require specific material handling.[16]

An untapped gold mine? Exploring the potential of market basket analysis to grow hotel revenue, International Market **Basket** Analysis identifies and predicts the purchasing behavior of customers based on the expenditure patterns of all previous customers. While widely applied in retail contexts, its use in hospitality is limited. This paper argues that Market Basket Analysis could increase revenue by enabling hotels to determine the most attractive additional products and services (beyond the room type) to offer new and repeat hotel guests. The method's potential is illustrated using five years of internal guest sales records from a luxury hotel group in Australia. Findings point to significant opportunities for hotel operators to use existing stored data to better understand purchasing decision patterns that can significantly increase revenue per transaction. Challenges to adoption and future research suggestions are offered[17]

A, MARKET BASE ANALYSIS IN DROPSHIP BUSINESS WITH APRIORI ALGORITHM IN DETERMINING R-BASED PRODUCT BUNDLING The union technique will link the statistics the use of a priori (experimental) rules that meet the minimum help necessities, i.E. The aggregate of each object in the database and the reliability necessities. Minimal confidence, that is, the power of the connection among the factors inside the association rule. In this look at, three-month income transaction statistics may be analyzed from dropshippers within the subject of style and beauty product income   using a combination method with a priori algorithm to look for patterns. Product grouping statistics. The minimal support values are the 10% and 20% self belief degrees, and with the % feature in which the regular shampoo items grow to be the primary made of the product institution, the 15 strongest product applications of fee are obtained. More desirable over 1.1. [18]

,Market basket analysis: Complementing association rules with minimum spanning trees This ponder proposes a strategy of advertise wicker container investigation based on negligible crossing trees, completing the seek for significant affiliation rules among a huge set of rules ordinarily particular to this investigation. . Much appreciated to the progressive tree structure of the hypermetric separations of the subdomains of the MST, the combined arrange makes it conceivable to discover solid interdependencies between items of the same sort and to discover items that play a part. is get to. get to or portal to a set of other closely related items. A significant viewpoint of this graph-based strategy is the ease with which item sets and bunches can take promoting activity. Applying our strategy to genuine value-based databases empowers you to: 1. uncover the interdependencies of higher esteem items, 2. uncover tall significance items with get to to another set of items, 3. characterize quality affiliation rules whereas identifying clusters and categorical connections among children's categories in general stores. Thi.. [19]

Influence-based approach to market basket analysis In this article, we propose an approach to market basket analysis based on the notion of social influence. While traditional market basket analysis looks for combinations of products that frequently co-occur in transactions, we seek to find a set of influential products that, if bought by a customer, will increase the sales volume of the shop. We believe that customers who purchase influential products would also be influenced to purchase other products. We validated our approach with two real-world datasets collected from online shoppings and one dataset collected from a supermarket concluding that influential products identified by our approach increase the influence spread with respect to different baselines: best-selling, highest centrality, frequent sequence initiator, and most promoted products.[20]

Penerapan Algoritma Apriori untuk Market Basket Analysis Data mining methods nowadays can offer assistance commerce proprietors to extend deals of their items. One well-known method is affiliation investigation. Affiliation investigation points to discover connections between things obtained

by clients. This kind of affiliation investigation is commonly known as advertise bushel investigation. Advertise bushel examination employments client information that has been put away on the base to

discover unused data in it. A priori calculation is an calculation for doing showcase ball investigation, which points to discover the things that are most regularly acquired. This a priori calculation produces an affiliation run the show that's advantageous for trade individuals. To select the most grounded affiliation run the show, it is fundamental to calculate the lift proportion. By calculating the elevator proportion of each affiliation run the show, you'll discover the substantial and most grounded affiliation rules. By conducting affiliation examination, it can be seen that client information can be utilized as input for commerce proprietors to decide deals procedures for their busness.[21]

Characteristics of Tourist Sources in Coastal Tourism Market Based on Cluster Examination At show, the commonly utilized strategies of highlight investigation of tourism sources in coastal tourism showcase with enormous information have destitute clustering degree of information highlight classification, so the inquire about on the characteristics of traveler sources in coastal tourism showcase based on cluster investigation is proposed. This paper partitions the visitor sources into diverse bunches with comparable traits or comparative relations. In selection of the objective work minimization methodology, the division and clustering of the traveler characteristics are completed. After it, the visitor source characteristic information is conducted by the closest neighbor engendering characteristic optimization clustering, so as to analyze the visitor source characteristics of the coastal tourism advertise. The recreation try is planned and compared with the clustering degree of highlight classification after enormous information preparing, demonstrating the investigate viability.[22]

a Market-Based Analysis on Small and Medium Business Strategies in Bogor This have a look at targets to fill the prevailing gap, particularly that the right education model may be advanced for MSME trainees so that they have a competitive advantage. Applying the MSME education model to sustainable Ciomas footwear can be completed in an integrated, targeted and targeted, marketplace-primarily based way. Using qualitative strategies, we recognise that the attributes of the production process, manufacturing device, production manage, buildings and centers, markets, exceptional standardization, commercial enterprise control Sales, finance and promotion have average overall performance. In this example, based on MSMEs perceptions of shoes, the performance degree is at the medium scale, while the expectations degree is on the excessive scale. [23]

A survey-based analysis of the academic job market, eLife Numerous postdoctoral analysts take after for college positions knowing moderately small approximately the enlisting way or what is needed to secure a errand offer. To manage with this need of understanding roughly the contracting way we performed a overview of candidates for staff positions: the study ran among May 2018 and May 2019, and procured 317 reactions. We analyzed the reactions to investigate the interaction among various academic measurements and enlisting impacts. We concluded that, over a positive edge, the benchmarks verifiably utilized to degree thinks about fulfillment – comprising of financing, number of distributions or diaries distributed in – had been incapable to totally separate candidates with and without work gives. Respondents too expressed that the enlisting strategy got to be superfluously stressing, time-ingesting, and lost in comments, notwithstanding of result. Our discoveries advocate that there's monster scope to improve the straightforwardness of the contracting handle.[24]

A new optimization model for market basket analysis with allocation considerations: A genetic algorithm solution approach , These days, advertise bushel investigation is one of the energizing inquire about ranges of data mining that has gotten more prominent consideration from analysts. But most extreme pertinent thinks about has focused on on conventional calculations and heuristic calculations with confining components not being the as it were affecting variables in wicker container advertise assessment. In this article to effectively show and look at the advertise bushel data, the optimization adaptation is proposed through considering the attribution parameter as one of the imperative and effective components of the wage proportion. . A hereditary calculation method is connected to resolve the issue of non-linear twofold programming, and a numerical example is utilized to demonstrate the advertised form. The results given appear that the arrangements procured appear more prominent sensible and significant. [24]
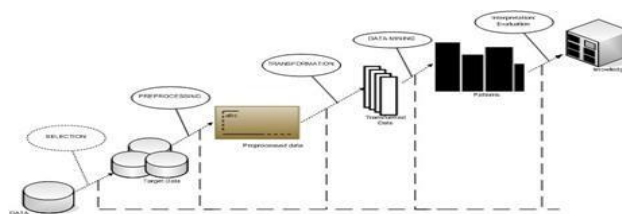
Fig. 1. Proses Knowledge Discovery in Database dan Data Mining

- Market Basket Analysis

Market Basket Analysis is a method or technique that is often used and the most useful for the marketing environment. The purpose of this Market Basket Analysis is to determine which products customers buy at the same time, where the name of this method is taken from the customer's habit of putting their goods in a basket or into a shopping list (market basket)[2]. Knowing which products to buy simultaneously can be of great help to traders or other companies. A store can also use this information to place frequently sold products together in an area or category.[7]

The advantage of using the Market Basket Analysis method is that besides being able to find out which products were purchased simultaneously, this method can use the information generated to re-order products for two or more products at the same time. This method can also be useful for top-level managers to be able to view purchase data from customers, so that they can know which ones are regular customers or who make the most purchases.

- Apriori Algorithm

Apriori algorithm is easy to execute and very simple, is used to mine all frequent itemsets in database. [8] The algorithm makes many searches in database to find frequent itemsets where kitemsets are used to generate k+1-itemsets. Each k- itemset must be greater than or equal to minimum support threshold to be frequency. Otherwise, it is called candidate itemsets. In the first, the algorithm scan database to find frequency of 1-itemsets that contains only one item by counting each item in database. The frequency of 1-itemsets is used to find the itemsets in 2- itemsets which in turn is used to find 3- itemsets and so on until there are not any more k-itemsets. International Journal on Natural Language Computing (IJNLC) Vol. 3, No.1, February 2014 23 If an itemset is not frequent, any large subset from it is also non-frequent [1]; this condition prune from search space in database.

- Association Rules

The availability of databases regarding the purchase transaction records of customers of a supermarket or another place, has led to the development of techniques that automatically find product associations or items stored in the database. An example is data about transactions in supermarkets. Transaction data lists all items purchased by a customer in a single purchase transaction. Managers want to know whether a group of items is always bought together. These managers can use this information to create supermarket layouts, so that the arrangement of these items can be optimal for each other or for promotional purposes, buyer segmentation, product cataloging, or viewing shopping patterns. The association rule wants to provide this information in the form of an "if-then" or "if-then" relationship computed from probabilistic data. [9] The idea of the association rule is to examine all possible if-then relationships between items and choose only the most likely ones as indicators of the relationship using the minimum support that the user has determined.

- Algoritma Eclat (Equivalence Class Transformation)

The Eclat algorithm performs frequent itemset searches from datasets. The Eclat algorithm introduced by Zaki, Parthasarathy, Ogihara, & Li (1997) is an algorithm that groups the same items

based on certain criteria into classes (Equivalence class). The same class is obtained from partitions of a set.   Equivalence

classes in the Eclat algorithm are built with prefix- based classes. The Eclat algorithm has a faster process, because the dataset will be presented in the vertical format of the dataset. [10]
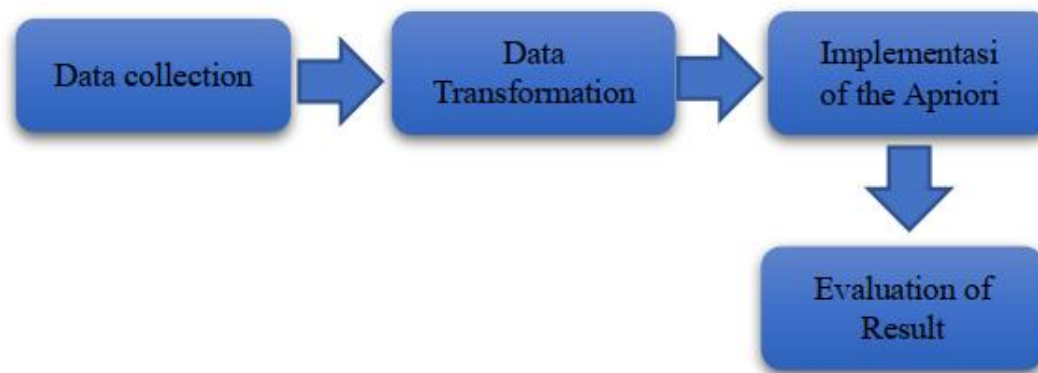
- Programming Language R

R (also known as GNU S) is a programming language and software for statistical and graphical analysis. The R language has now become the de facto standard among statisticians for statistical software development, and is widely used for statistical software development and data analysis.

2. Method

In this research activity we divide it into 4 stages, namely:

Fig. 2. Tahapan penelitian

Beginning with data collection, the dataset uses clothes shop data with interview techniques. Then data transformation is used to process data before the data mining process stage performs R

programming. Then the a priori algorithm implementation at this stage uses Market Basket Analysis with a priori algorithm using R programming language. The last one evaluating the results is evaluating and checking the results to match what is specified.

2.1. Data collection

In data collection, we use the direct interview technique to the clothes shop owner because the clothes shop does not have a database system because the sales are still manual, even though in one marketplace the data system is still integrated into the marketplace system. From the interview, there are 3,450 transaction data and 69 items which we are ready to process into a dataset in TSV format.

Fig. 3.clothes shop sales

2.2. Data Transformation

In the data transformation stage, we create a data frame to facilitate data processing in R programming, the steps are as follows: It can be saved in .tsv format and then converted to data.frame to make it easier to process data in R language.

With the instructions in the R programming language, you can:

- Library (arules) Library (tidyverse)

- transactionsead.transactions (file = "C:

- /Users/isuta/Downloads/transaksi_dqlab_retail.tsv", format = "single", sep = "\ t", cols = c (1,2), skip = 1)

- print (transaction)

Fig. 4.read transaction files to R programming

From processing the data it is known: 3450 transactions with 69 items. Frame transaction data

- data_item <- itemFrequency (transaction, type = "absolute")

- data_item <- sort (data_item, decreasing = TRUE)

- data_item <- data.frame ("Product Name" = names (data_item), "Amount" =

- data_item, row.names = NULL)

- print   (data_item).

2.3. **Implementation of the Apriori Algorithm**

Technically, the a priori algorithm will look for the level of association between items in many combinations of data groups automatically. This combination can also be arranged with an association rule "If you buy this product A, you will buy product B", so this algorithm is categorized as Association Rules in the realm of machine learning. The transformed dataset is then read with Apriori's instructions which are available in Programming R by determining the minimum support, confident and lift. The following is a datamining design flowchart that is used as follows:
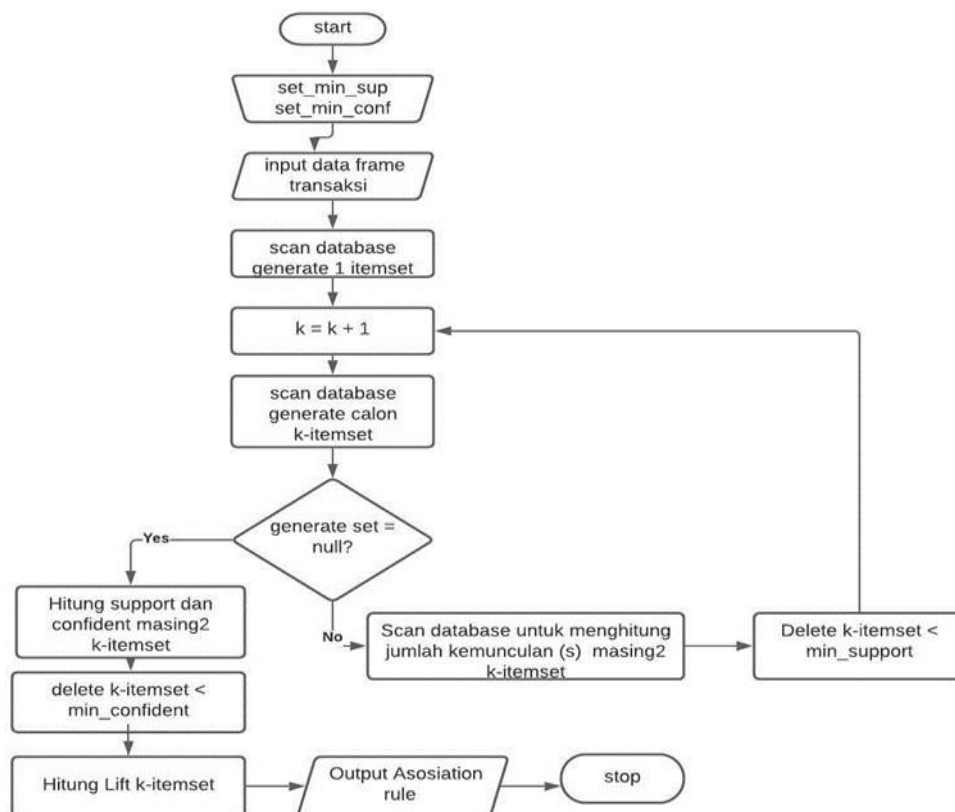
Fig. 5. Flowchart Diagram

3. Results and Discussion

The transformed dataset is then read with Apriori's instructions which are available in Programming R by determining the minimum support, confident and lift. Our first step is to display a matrix of transactions with the command in R to see the distribution of items across transactions with the code program.

transaction <-read.transactions(file = "C:

/Users/isuta/Downloads/transaksi_dqlab_retail.tsv", format = "single", sep = "\ t", cols = c (1,2), skip = 1)

- transaction @ data
- itemFrequencyPlot (transaction)
- Item Distribution Matrix.
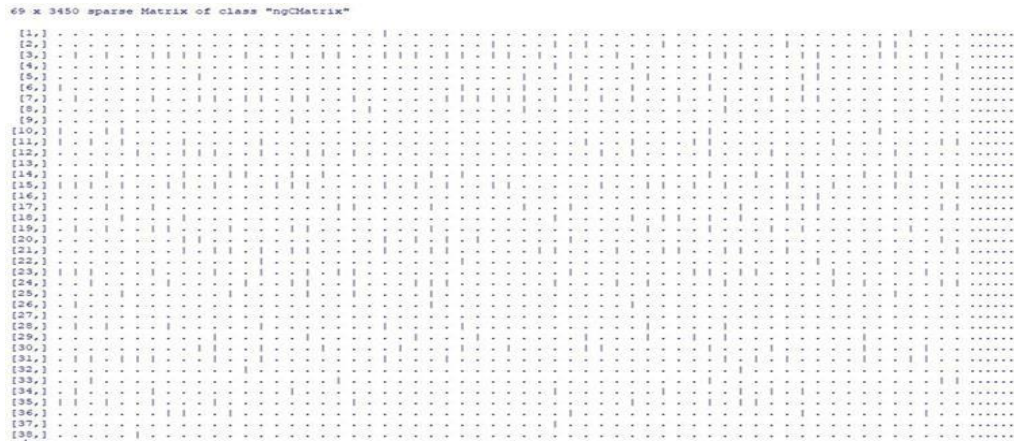
Explanation of the results:

- 69 X 3450 means that there are 69 items and 3450 transactions

- [1,] and [3450,] represent index of each item in transaction transaction

- Period symbol. indicates the item is not in the transaction.

- Straight line symbol | indicates that the item is in the transaction.

From the graph and metrix, we can analyze the associations between products with a priori algorithms that already exist in R language programming by determining the minimum support and confidence.

Here we define the parameters:

- Support: 0.1 which is a 10% occurrence rate.

- Confident: 0.5 which is 50% of the support value for all transactions.

- transaction <-read.transactions(file="C:

- /Users/isuta/Downloads/transaksi_dqlab_retail.tsv", format = "single", sep = "\ t", cols = c (1,2), skip = 1)

- > mba <- a priori (transactions, parameters = list (supp = 0.1, confidence = 0.5)).

From the results of the a priori algorithm, there are 53 association rules, of the 53 we inspect them to look for strong associations.

- Inspect rule associations.

```
set of 53 rules
> inspect(mba)
     lhs                                                      rhs                 support   confidence coverage  lift      count
[1]  {}                                                    => {Shampo Biasa}      0.6014493 0.6014493 1.0000000 1.0000000 2075
[2]  {Tali Pinggang Anak}                                  => {Serum Vitamin}     0.1037681 0.6884615 0.1507246 1.4096097 358
[3]  {Tas Waist Bag}                                       => {Serum Vitamin}     0.1298551 0.8057554 0.1611594 1.6497662 448
[4]  {Tas Ransel Mini}                                     => {Shampo Biasa}      0.1037681 0.6392857 0.1623188 1.0629088 358
[5]  {Tank Top}                                            => {Shampo Biasa}      0.1002899 0.5844595 0.1715942 0.9717519 346
[6]  {Celana Tactical }                                    => {Shampo Biasa}      0.1156522 0.6605960 0.1750725 1.0983404 399
[7]  {Sweater Top Panjang}                                 => {Shampo Biasa}      0.1026087 0.5880399 0.1744928 0.9777048 354
[8]  {Flat Shoes Ballerina}                                => {Shampo Biasa}      0.1159420 0.6201550 0.1869565 1.0311011 400
[9]  {Celana Pendek Casual}                                => {Shampo Biasa}      0.1127536 0.6087637 0.1852174 1.0121613 389
[10] {Blouse Denim}                                        => {Serum Vitamin}     0.1072464 0.5882353 0.1823188 1.2043987 370
[11] {Blouse Denim}                                        => {Shampo Biasa}      0.1292754 0.7090620 0.1823188 1.1789224 446
[12] {Hair and Scalp}                                      => {Baju Kemeja Putih} 0.1078261 0.5329513 0.2023188 1.4650852 372
[13] {Hair and Scalp}                                      => {Serum Vitamin}     0.1026087 0.5071633 0.2023188 1.0384056 354
[14] {Hair and Scalp}                                      => {Shampo Biasa}      0.1136232 0.5616046 0.2023188 0.9337522 392
[15] {Dompet Card Holder}                                  => {Shampo Biasa}      0.1449275 0.6476694 0.2237681 1.0769462 500
[16] {Tas Travel}                                          => {Shampo Biasa}      0.1168116 0.5289714 0.2208696 0.8793283 403
[17] {Dompet Flip Cover}                                   => {Serum Vitamin}     0.1165217 0.5121019 0.2275362 1.0495173 402
[18] {Dompet Flip Cover}                                   => {Shampo Biasa}      0.1495652 0.6573248 0.2275362 1.0929015 516
[19] {Celana Pendek Jeans}                                 => {Shampo Biasa}      0.1466667 0.6239211 0.2350725 1.0373628 506
[20] {Hair Dryer}                                          => {Serum Vitamin}     0.1226087 0.5152253 0.2379710 1.0549124 423
[21] {Hair Dryer}                                          => {Shampo Biasa}      0.1481159 0.6224117 0.2379710 1.0348532 511
[22] {Celana Jeans Sobek Wanita}                           => {Shampo Biasa}      0.2060870 0.8525180 0.2417391 1.4174395 711
[23] {Wedges Hitam}                                        => {Shampo Biasa}      0.1460870 0.5936396 0.2460870 0.9870152 504
[24] {Sepatu Sport merk Z}                                 => {Serum Vitamin}     0.1414493 0.5495495 0.2573913 1.1251905 488
[25] {Sepatu Sport merk Z}                                 => {Shampo Biasa}      0.1536232 0.5968468 0.2573913 0.9923478 530
[26] {Tali Pinggang Gesper Pria}                           => {Serum Vitamin}     0.1481159 0.5094716 0.2907246 1.0431317 511
[27] {Tali Pinggang Gesper Pria}                           => {Shampo Biasa}      0.1913043 0.6580259 0.2907246 1.0940672 660
[28] {Sepatu Sandal Anak}                                  => {Serum Vitamin}     0.2150725 0.6986817 0.3078261 1.4305353 742
[29] {Sepatu Sandal Anak}                                  => {Shampo Biasa}      0.2237681 0.7269303 0.3078261 1.2086311 772
[30] {Cover Koper}                                         => {Shampo Biasa}      0.1959420 0.6224678 0.3147826 1.0349464 676
[31] {Celana Jogger Casual}                                => {Shampo Biasa}      0.2084058 0.6329225 0.3292754 1.0523290 719
[32] {Baju Kemeja Putih}                                   => {Serum Vitamin}     0.1866667 0.5131474 0.3637681 1.0506579 644
[33] {Baju Kemeja Putih}                                   => {Shampo Biasa}      0.2437681 0.6701195 0.3637681 1.1141746 841
[34] {Baju Batik Wanita}                                   => {Serum Vitamin}     0.2049478 0.5373476 0.3802899 1.1002072 705
[35] {Baju Batik Wanita}                                   => {Shampo Biasa}      0.2463768 0.6478659 0.3802899 1.0771746 850
[36] {Serum Vitamin}                                       => {Shampo Biasa}      0.3162319 0.6474777 0.4884058 1.0765293 1091
[37] {Shampo Biasa}                                        => {Serum Vitamin}     0.3162319 0.5257831 0.6014493 1.0765293 1091
[38] {Celana Jeans Sobek Wanita,Tali Pinggang Gesper Pria} => {Shampo Biasa}      0.1052174 0.8897059 0.1182609 1.4792700 363
```

- Of the 53 rules, we will re-sort them into 5 rules which are closely related between products by sorting the amount of support

- Code Program ;

- inspect(sort(rules, by = 'support')[1:5])

```
> inspect(sort(rules, by = 'support')[1:5])
    lhs                 rhs             support   confidence coverage  lift     count
[1] {}                => {Shampo Biasa}  0.6014493 0.6014493 1.0000000 1.000000 2075
[2] {Serum Vitamin}   => {Shampo Biasa}  0.3162319 0.6474777 0.4884058 1.076529 1091
[3] {Shampo Biasa}    => {Serum Vitamin} 0.3162319 0.5257831 0.6014493 1.076529 1091
[4] {Baju Batik Wanita} => {Shampo Biasa} 0.2463768 0.6478659 0.3802899 1.077175 850
[5] {Baju Kemeja Putih} => {Shampo Biasa} 0.2437681 0.6701195 0.3637681 1.114175 841
>
```

Before we conclude the strongest association, we use a comparison algorithm, the Eclat Algorithm. The results of the Eclat Algorithm Association in the R programming language with a minimum support of 0.1 are the same as the a priori algorithm, namely.

Code program : rules = eclat(data = transaksi, parameter = list(support = 0.1, minlen = 2)).

```
> rules = eclat(data = transaksi, parameter = list(support = 0.1,  minlen
Eclat

parameter specification:
 tidLists support minlen maxlen              target   ext
    FALSE     0.1     2     10 frequent itemsets TRUE

algorithmic control:
 sparse sort verbose
      7   -2    TRUE

Absolute minimum support count: 345

create itemset ...
set transactions ...[69 item(s), 3450 transaction(s)] done [0.01s].
sorting and recoding items ... [40 item(s)] done [0.00s].
creating bit matrix ... [40 row(s), 3450 column(s)] done [0.00s].
writing  ... [63 set(s)] done [0.00s].
Creating S4 object  ... done [0.00s].
```

From the results of the Eclat Algorithm, there are more rules, namely 63 association rules. Here we take 5 higher support values, namely

```
> inspect(sort(rules, by = 'support')[1:10])
       items                                    support
 [1]   {Serum Vitamin,Shampo Biasa}             0.3162319
 [2]   {Baju Batik Wanita,Shampo Biasa}         0.2463768
 [3]   {Baju Kemeja Putih,Shampo Biasa}         0.2437681
 [4]   {Sepatu Sandal Anak,Shampo Biasa}        0.2237681
 [5]   {Sepatu Sandal Anak,Serum Vitamin}       0.2150725
```

The results of the eclat algorithm are almost the same as the priori algorithm. However, the lack of the Eclat algorithm cannot determine the values of confident and lift.

3.1. Evaluation of Results

In the final step, we evaluate the results of processing the data so that it becomes the following information:

```
> inspect(sort(rules, by = 'support')[1:5])
    lhs                     rhs              support   confidence coverage  lift     count
[1] {}                   => {Shampo Biasa}   0.6014493 0.6014493  1.0000000 1.000000 2075
[2] {Serum Vitamin}      => {Shampo Biasa}   0.3162319 0.6474777  0.4884058 1.076529 1091
[3] {Shampo Biasa}       => {Serum Vitamin}  0.3162319 0.5257831  0.6014493 1.076529 1091
[4] {Baju Batik Wanita}  => {Shampo Biasa}   0.2463768 0.6478659  0.3802899 1.077175  850
[5] {Baju Kemeja Putih}  => {Shampo Biasa}   0.2437681 0.6701195  0.3637681 1.114175  841
> |
```

- Know the 5 association rules that can be used as information.

- Know 2 strong association rules and can be recommended for information on promotional items.

- Recommended special package products, namely regular shampoo with vitamin serum.

- Display on the ecomerse by recommending the usual sahampo with women's batik clothes.

- We can package products that are not selling well with items that are selling well so that we can control inventory.

- Know the pattern of product combinations from sales data with a lot of available information.

4. Conclusion

In this paper, based on the results of this study it can be concluded that:

- Know the association rules that can be used as information.

- Comparing other algorithms with a priori algorithms.

- Contains special package product recommendations.

- Can be displayed on the ecomerse by recommending goods that have a strong association.

- We can package products that are not selling well with items that are selling well so that we can control inventory

References

[1] L. Maulida, "Application of Datamining in Grouping Tourist Visits to Top Attractions in Prov. Dki Jakarta With K-Means," JISKA (Jurnal Inform. Sunan Kalijaga), vol. 2, no. 3, p. 167, 2018, doi: 10.14421/jiska.2018.23-06.

[2] A. F. Yudanar, S. H. Fitriasih, and M. Hasbi, "Recommendations for Goods In Electrical Stores Using A priori Algorithms," J. Teknol. Inf. and Komun., vol. 8, no. 2, 2020, doi: 10.30646/tikomsin.v8i2.499.

[3] Y. Akbar and M. J. Hakim, "Market Basket Analysis Using Language-Based A priori Algorithm R (Transretail Indonesia Case Study)," CKI SPOT, vol. 11, no. 2, pp. 173–180, 2018.

[4] I. Qoniah and A. T. Priandika, "Basketball Market Analysis To Determine Rule Association With A priori Algorithm (Case Study: Tb. Menara)," J. Teknol. and Sist. Inf., vol. 1, no. 2, pp. 26–33, 2020.

[5] T. Brian and A. Sanwidi, "Implementation of A priori Algorithms for Market Basket Analysis Based on R," J. ELTIKOM, vol. 2, no. 1, pp. 1–8, 2018, doi: 10.31961/eltikom.v2i1.40.

[6] I. A. Nikmatun and I. Waspada, "Implementation of Data Mining for Classification of Student Study Period Using K-Nearest Neighbor Algorithm," J. SYMMETRY, vol. 10, no. 2, pp. 421–432, 2019.

[7] R. Bakri, A. Halim, and N. P. Astuti, "Product Marketing Strategy Information System with Market Basket Analysis and Sales Forecasting Method: Makassar City Supermarket," J. Manaj. Theor. and Terap. | J. Theory Appl. Manag., vol. 11, no. 2, p. 89, 2018, doi: 10.20473/jmtt.v11i2.9769.

[8] U. Ciputra, U. J. Soedirman, and L. P. P. W. Parahita, "Association Role Mining on Moodle using A priori Algorithm: Case Study of UN IPA Values," Semin. Nas. Aptikom..., pp. 355–366, 2019, [Online]. Available: http://publikasi.dinus.ac.id/index.php/semnastik/article/view/2961.

[9] H. R.M. Tambun and A. Anofrizen, "Design a Data Mining Application To Display Student Graduation Rate Information With A priori Algorithm," J. Ilm. Engineering and Manaj. Sist. Inf., vol. 1, no. 1, pp. 31–38, 2015, [Online]. Available: http://ejournal.uin-suska.ac.id/index.php/RMSI/article/view/1302/pdf_3.

[10] [C.Pradeepkumar and S.Loganathan, "Application of Association Methods Using A priori Algorithms In Consumer Shopping Pattern Applications (Gramedia Bintaro Bookstore Case Study)," Int. J. Sci. Eng. Res. (IJ0SER), vol. 3, no. 4, p. 2, 2015, [Online]. Available: http://journal.uinjkt.ac.id/index.php/ti/article/view/5602/3619.

[11] Yüksel Akay Ünvan, Market basket analysis with association rules, Communications in Statistics - Theory and Methods (2021)

[12] [Srinivasa Kumar VRenganathan RVijayaBanu C et al.Consumer buying pattern analysis using apriori association rule, , International Journal of Pure and Applied Mathematics, (2018), 119(Special Issue 7C)

[13] Heydary, M. and Yousefli, A. (2017), "A new optimization model for market basket analysis with allocation considerations: A genetic algorithm solution approach", Management & Marketing. Challenges for the Knowledge Society, Vol. 12, No. 1, pp. 1-11, DOI: 10.1515/mmcks-2017-0001

[14] PView M et al ,A Study on Market Basket Analysis and Association Mining,In Proceedings of National Conference on Machine Learning, (2019), (April)

[15] Halim SOctavia TAlianto C, Designing facility layout of an amusement arcade using market basket analysis, Procedia Computer Science, (2019)

[16] Solnet DBoztug YDolnicar S,An untapped gold mine? Exploring the potential of market basket analysis to grow hotel revenue, International Journal of Hospitality Management, (2016), 56

[17] Fauziyyah A, MARKET BASE ANALYSIS IN DROPSHIP BUSINESS WITH APRIORI ALGORITHM IN DETERMINING R-BASED PRODUCT BUNDLING,Indonesian Journal of Business Intelligence (IJUBI), (2019), 2(1)

[18] Valle MRuz GMorrás R,Market basket analysis: Complementing association rules with minimum spanning trees

Valle MRuz GMorrás R, Expert Systems with Applications, (2018), 97

[19] Monteserin AArmentano M, Influence-based approach to market basket analysis, Information Systems, (2018), 78

[20] FAHRUDIN N, Penerapan Algoritma Apriori untuk Market Basket Analysis, MIND Journal, (2019), 1(2)

[21] Wang X et al , Characteristics of Tourist Sources in Coastal Tourism Market Based on Cluster Analysis

, Journal of Coastal Research, (2020), 103(sp1)

[22] Gen Gen Gendalasari, a Market-Based Analysis on Small and Medium Business Strategies in Bogor ' S Footwear Industry, Jurnal Aplikasi ekonomi Akuntansi dan Bisnis, (2020), 2(1)

[23] Fernandes JSarabipour SSmith C et al.See more, A survey-based analysis of the academic job market, eLife, (2020), 9

[24] Heydari MYousefli A, A new optimization model for market basket analysis with allocation considerations: A genetic algorithm solution approach, Management and Marketing, (2017), 12(1)