Market Basket Analysis For Procurement Of Food Stock Using Apriori Algorithm And Economic Order Quantity

Made Sri Indradewi Adnyana[1], Rukmi Sari Hartati[2], Made Sudarma[3]

[1]Department of Electrical and Computer Engineering, Post Graduate Program, Udayana University
[2,3] Department of Electrical and Computer Engineering, Udayana University
email: [1]indradewi2685@gmail.com, [2]rukmisari@unud.ac.id, [3]msudarma@unud.ac.id

Abstract — In a company, it's important to think about inventory patterns. Because without inventory management, companies can suffer losses. Perceived losses if the goods owned in the company's stock are lacking but the interest in goods is very high can cause consumers to switch to other companies and will reconsider making transactions. This also happens if the goods available in the warehouse have a large amount but the consumer's interest in the goods will experience a slight loss especially if the item has an active period so there is a buildup or loss. Therefore an item management strategy is needed. One company that can use both of these methods is Starbucks Legian. Because using the Economy Order Quality method can determine the number of subsequent orders and using a priori algorithm can help companies determine the items that need to be held more and can be given as promos with other products sold or given discounts. Using these two methods will attract consumers to shop at Starbucks Legian.

Index of Terms - Algorithm Apriori, Data Mining, EOQ, Inventory.

I. INTRODUCTION

In a company important marketing strategy. In terms of marketing strategy. Because receiving goods is a very important thing to consider before deciding on a marketing strategy to be carried out. Availability of goods is very important because the management can determine the amount of goods stored in the warehouse to avoid goods entering the minimum quantity and excessive amount of goods. Because in some companies the goods sold during their usefulness. This can cause losses on goods issued in excess of the amount to be paid if the minimum is needed, then consumers will turn to other companies and will reconsider making transactions for the same goods at companies that have a minimum amount.

One technique that is carried out in a company for inventory management and marketing strategies. Stock management can use data mining to help companies make decisions in their business. After managing stock of goods, a business strategy that can be carried out is by also being able to transfer goods according to the number of transactions that are the most of the most desirable items. Stock inventory management can be done using market basket analysis techniques,

Market basket analysis is used to determine the relationship between products and consumers. One example is owned by market basket analysis using association rules. One algorithm that uses association rules is a priori algorithm. A priori algorithm is used to give advice by doing each product in a transaction. After making the previous steps in the company's business strategy, it is also necessary to consider the number of subsequent orders for the goods that will be most desirable, which will be given recommendations. Economical order quantities model can also be used to help determine the number of future orders and can find out the number of further orders.

One company that can use a priori algorithms and economical order quantity methods to carry out marketing strategies is Starbucks legian. Starbucks company is a company that is famous for its coffee. In addition to Starbucks coffee also has some food that can be used as a companion to enjoy coffee or can be used as a snack. Food provided can also be given as a promo for each transaction carried out. By using these two techniques to help Starbucks Legian in business strategies in marketing food that is in accordance with the interests of consumers.

II. LITERATURE REVIEW

A. Data Mining

Data mining according to research conducted by [1] is a technique in data processing to determine meaningful relationships using statistical techniques, mathematics, and machine learning. . According to research conducted by [2] Data mining is a term that can be used to describe the process of extracting value / information from a database. Data mining has the empathy needed to mine data effectively.

According to an explanation of data mining in research conducted by [3] explains data mining has several stages that are carried out with the initial stages by selecting the data taken

from the source data to the target, then the pre-processing stages are used to improve data quality, data mining, data transformation and the final stages namely interpretation and evaluation that can produce an output that can be in the form of new knowledge that is expected to provide a better contribution.

According to research conducted by [4] data mining is an analytical tool that allows users to analyse data from different dimensions, categorize data and summarize relationships with data.

B. Apriori Algorithm

According to research conducted by [5] apriori algorithm is a very well-known algorithm using the association rule technique in searching frequent itemset. According to research also conducted by [6] apriori algorithm is also a process of data retrieval to determine the associative relationship of a combination of items in an associative relationship. In its use a priori algorithm uses knowledge of frequent itemset that has been known before to process further information. According to research conducted by [7] using a priori algorithm suggests the relationship between several attributes so that it is often referred to as affinity analysis or market basket analysis Apriori algorithm is used to determine candidates who might emerge with regard to minimum support. In making calculations using the association rule technique uses the formula to find support and confidence. According to research conducted by [8] explains the notion of :

a. *Support* is a probability owned by consumers in buying several products simultaneously from all transactions carried out. The size obtained will be a determination whether an item / itemset is feasible to find its confidence value. Usually done in the initial stages of analysing high frequency patterns in the form of mathematical equations of the support values as follows:

$$Support(X=>Y) = P(X \cap Y)$$

Information :

X => Y is an item that will appear together

P (X∩Y is a transaction probability containing the values of X and Y divided by the total number of transactions made.

b. *Confidence* is the probability of an occurrence of several products purchased simultaneously where one product is definitely purchased. Confidence becomes the second stage with the mathematical formula, namely:

$$Confidence(X=>Y) = P(Y \mid X)$$

Information :

P (X | Y) = P (Y | X) are items that appear together

X∩Y is a transaction probability containing X and Y divided by transactions containing the value X.

According to the research conducted by [9] in this study explaining the way the a priori algorithm works that the a priori algorithm will produce new candidates from the k-itemset of the itemset that often appeared in the previous step and calculate the value of itemset support.

C. Economic Order Quality (EOQ)

According to research conducted by [10] economic order quantity method is a method of inventory management that is classic and simple so that it is used to calculate the total inventory cost minimization based on the equilibrium level or the point of equilibrium of the cost and store cost curve. Next step is the reoder point step. Reorder point is the specific time a company has to place a product order again, so that the right time comes with the end of the product purchased, especially by using the EOQ method. Then the calculation process determines the maximum inventory (Maximum inventory). This step is used in companies to determine the amount of inventory in the warehouse so that it is not excessive to avoid working capital breakthroughs. According to research conducted by [11] explained to calculate the EOQ model can be done in three ways namely:

1. Using a table (tabular)

   Ensure the amount of an economic order by compiling a list (table) of the number of costs per year as well as the number of orders. This happens where the EOQ value is determined when the Carrying Cost is the same as the Ordering Cost.

2. Using graphs (Graphical)

   Ensure the number of economic orders by drawing a Caring Cost chart, an Ordering Cost chart and a Total Cost Graph. All three graphs are made in one picture where the vertical axis proves the costs and the horizontal axis proves the number of orders for one year.

3. Using the formula (Formula)

   Ensuring the number of economic orders by involving into a predetermined mathematical formula.

III. METODOLOGY

3.1 Data Collection Methods

In this study, there are 3 data collection methods. The techniques used in this study are 3, namely:

• Interview Techniques

Interviews were conducted at Starbucks Legian using the semi-structured interview technique method. According to research conducted by [12] the semi-structured technique interview method becomes useful when the online data collection process..

• Literature Study

Literature study in this study discusses several journals that discuss procurement of goods, a priori algorithms, and the quality of economic order. Studio data literature is used as a basis for making this journal

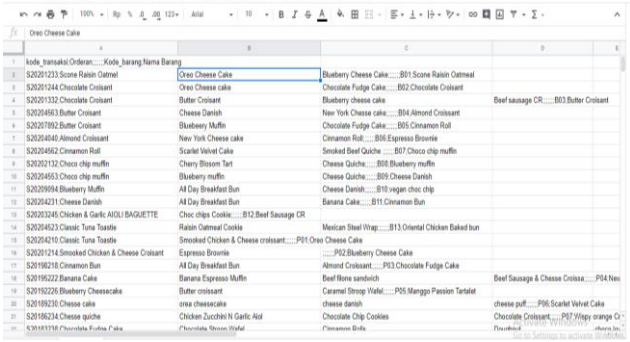
3.2 Method of Analysis

The analytical method contained in this study is the data collected during the data search process and collected which will be carried out towards data analysis. The method of analysis here is distinguished by its type, types and types of primary and secondary data. Primary data in this study are data on food sales at Starbucks Legian stores and list of goods order

data. Secondary data is the SOP for receiving and ordering goods

IV. RESULTS AND DISCUSSION

Based on the identification of problems that must be solved first, namely by making initial preparations from the data to be processed. Furthermore, the data obtained are 500 times the transactions made during the process..



The next step to do is to change the code. The language used in this study uses scikit-learn in python..

4.1 Implements Apriori Algorithm

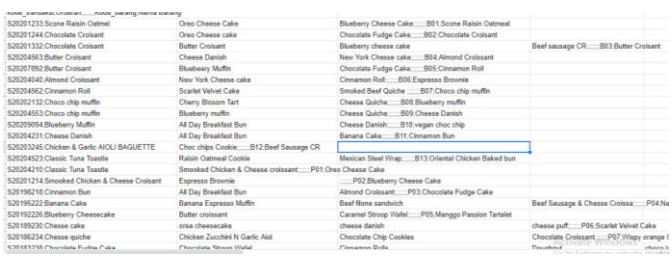
The next step to take is to prepare a 1-itemset candidate list as the table below.

TABLE 1
1-ITEMSET CANDIDATE

kandidat item-1 set	item set
CINNAMON BUN	25
ALL DAY BREAKFAST BUN	71
Almond Croissant	60
BALI 64% CHOCOLATE	6
Banana Cake	20
BANANA ESPRESSO MUFFIN	4
Beef Filone Sandwich	3
Beef Sausage & Cheese Croissa	17
BLUEBERRY CHEESECAKE	56
Butter Croissant	87
Caramel Stroop Wafel	16
CHEESE CAKE, Oreo Cheesecake,	59
CHEESE DANISH	44
Cheese Puff	1
Cheese Quiche	67
CHICKEN ZUCCHINI N GARLIC AIOL	25
Chocolate Chip Cookies	33
Chocolate Croissant	60
CHOCOLATE FUDGE CAKE	61
Chocolate StroopA23:A27 Wafel	11
Cinnamon Rolls	69
DOUGHNUT, CHOCO LOVE DONUT, R	8
Espresso Brownies	55

After determining the 1-item candidate, the next step is made into a 2-item candidate, the following results will be obtained:

TABLE II
LIST OF CANDIDATE 2-ITEMSET



The next step is to calculate the value of support and confidence of each item often so that it will make the association rules. This step will be done on scikit-learn by using Python language, as in the picture below

```
Support(B) = (Transactions containing (B))/(Total Transactions)
```

Fig.1 *Support formula*

After finding the support value, determine the value of trust that can be applied with the formula below:

```
Confidence(A→B) = (Transactions containing both (A and B))/(Transactions containing A)
```

Fig.2 *Confidence* formula

Next determine the value of the elevator (A → B). This value is used based on the ratio of sales increasing B compilation A sold. Lift value can be calculated by dividing the value of confidence divided by the value of support. The mathematical formula used is as below:

```
Lift(A→B) = (Confidence (A→B))/(Support (B))
```

Fig.3 Lift formula

The function of the elevator value is used to provide recommendations for goods that are likely to be purchased simultaneously or the sale value will be the same as the highest value.

After all the values of support, confidence and lift values are found next to implement the Python programming language.

```
for i in range(0, num_record):
    records.append([str(df.values[i, j])
                    for j in range(0, 20)])

association_rules = apriori(records, min_support=0.015, min_confidence=0.2, min_lift=3, min_length=2)
association_results = list(association_rules)

print(len(association_results))
print(association_results[0])

results = []
for item in association_results:
    pair = item[0]
    items = [x for x in pair]

    value0 = str(items[0])
    value1 = str(items[1])
    value2 = str(item[1])[:7]
    value3 = str(item[2][0][2])[:7]
    value4 = str(item[2][0][2])[:7]

    row = (value0, value1, value2, value3, value4)

    results.append(row)

    label = ['Title1', 'Title2', 'Support', 'Confidence', 'Lift']
```

Fig 4. code phyton

After the code is run it will look like the Fig below::

```
126
RelationRecord(items=frozenset({'Smoked Beef Quiche', ' '}), support=0.0270270
              Title1 Title2  Support Confidence     Lift
0  Smoked Beef Quiche         0.02702    0.30000  0.30000
               Title1 Title2  Support Confidence     Lift
0   Smoked Beef Quiche         0.02702    0.30000  0.30000
1  classic tuna toasti         0.02702    0.30000  0.30000
```

Fig 5. Implements Algorithm Apriori

After getting 2 types of foods that are recommended can be done using the EOQ method.

4.2 Economic Order Quality (EOQ)

At this stage it is used to calculate EOQ for each food that is most recommended based on apriori algorithm.

TABLE III
TABLE EOQ FOOD RECOMMENDATIONS

Nama Barang	Qty	Harga	Harga Total	Biaya Pesan atau (10%)
Smooked Beef Quiche	46	Rp 43.000	Rp 1.978.000	Rp 197.800
Classic Tuna Toasti	44	Rp 48.000	Rp 2.112.000	Rp 211.200

1. Calculate EOQ Smoked Beef Quiche

   $EOQ = \frac{\sqrt{2xDxS}}{H}$

   $EOQ = \frac{\sqrt{2x46x197800}}{43000x10\%} =$

   $EOQ = \frac{\sqrt{2x46x197800}}{4300} = \sqrt{4232} = 65{,}05$

   So, the EOQ for smoked beef quiche is 65.05 or 65 pieces
2. Calculate EOQ Classic Tuna Toastie

   $EOQ = \frac{\sqrt{2xDxS}}{H}$

   $EOQ = \frac{\sqrt{2x44x211200}}{48000x10\%} =$

   $EOQ = \frac{\sqrt{2x44x211200}}{4800} = \sqrt{3872} = 62{,}22$

   So, EOQ for classic tuna toastie foods is 62.22 or 62 pieces.

After performing the calculation steps of the most recommended items using a priori algorithm. Next through the Reorder Point step. This stage is done by the formula:

Reorder point = use during lead time + safety stock.

Use during lead time = lead time x material users per day. If Starbucks has 5 days with an average safety stock of 10 pieces for each product, the following calculation will occur :

1. ROP Smoked Beef Quiche
   Average sales a day = 65/28 = 2.32
   Reorder Points = (5 * 2.32) + 10 = 21.6 pieces rounded up to 22 pieces
2. ROP Classic Tuna Toastie
   Average sales a day = 62/28 = 2.21
   Reorder Points = (5 * 2.21) + 10 = 21.05 pieces rounded up to 21 pieces

After the Reorder point stage, the final step will be to determine the maximum supply. Determination of food inventory with the formula maximum inventory = Safety stock + EOQ.

1. *Maximum Inventory Smoked Beef Quiche*
   Maximum inventory = 10 + 65 = 75 pieces
2. *Maximum Inventory Classic Tuna Toastie*
   Maximum inventory = 10 + 62 = 72 pieces

V. CONCLUSION

5.1 CONCLUSION

Based on the research that has been done, it can be concluded that using these two methods can determine the pattern of food inventory so that it can regulate business strategies. Based on the a priori algorithm method, 2 foods are obtained from transaction data obtained in February. The results can help provide food recommendations for promos such as smoked beef quiche and Classic Tuna Toastie. Whereas using the EOQ method can help Starbucks Legian determine the pattern of goods needed for the following month and can determine the maximum number of orders. Integration can help Starbucks Legian reduce losses..

5.2 RECOMMENDATION

Based on the research conducted, it was obtained a suggestion. The EOQ method can then be continued in the program so that the program becomes more perfect. Future studies can also develop further research by adding several supporting methods.

VI. ACKNOWLEDGEMENT

Thank you, Starbucks Legian, for having to assist the researchers in the research analysis of data mining for food supplies at Starbuck Legian. With the support of the Head of Shop and all Starbucks staff, this research is complete. Thank you for all of your input and feedback for this research

REFERENCES

[1] S. Wahyuni, Suherman and P. L. Harahap, "Implementasi Data Mining Dalam Mempredik Stok Barang Menggunakan Algoritma Apriori," *Jurnal Teknik dan Informatika*, vol. 5, no. 2, pp. 67-71, 2018.

[2] H. P. B. Surya, R. L. Rahardian and M. Sudarma, "Application of Neutral Network Overview In Data Mining," *International Journal of Engineering and Emerging Technology*, vol. 2, no. 1, pp. 94-96, 2017.

[3] K. Fatmawati and A. P. Windarto, "Data Mining Penerapan Rapidminer dengan K-Means Cluster Pada Daerah Terjangkit demam berdarah dengue (DBD) berdasarkan provinsi," *Jurnal of Computer Engineering System and Science*, vol. 3, no. 2, pp. 173-178, 2018.

[4] I. G. N. A. Mahendra, I. B. L. M. Suta and M. Sudarma, "Classification of Data Mining with Adaboost Method in Determining Credit Providing for Customers," *International Journal of Engineering and Emerging Technology*, vol. 4, no. 1, pp. 31-36, 2019.

[5] D. Lingga, "Penerapan Algoritma Apriori Dalam Memprediksi Persediaan Buku Pada Perpustakaan SMA Dwi Tunggal Tangjung Morawa," *Majalah Ilmiah Informasi dan teknologi Ilmiah*, vol. 4, no. 1, pp. 18-22, 2016.

[6] D. A. K. Arimbawa P, I. N. A. Prabawa and P. A. Mertasana, "Implementation of Apriori Algorithm in Determining Tourism Visit Patterns To Bali," *Internasional Journal of Engineering And Emerging Technology*, vol. 4, no. 1, pp. 10-14, 2019.

[7] A. P. S. Iskandar, K. A. Shobirin and K. O. Saputra, "Analysis of Shopping Cart At Drugs Store By Using An

Apriori Algorithm," *International Journal Of Engineering And Emerging Technology*, vol. 2, no. 1, pp. 97-103, 2017.

[8] O. Pahlevi and A. Sugandi, "Penerapan Algoritma Apriori dalam Pengendalian Kualitas Produk," *Jurnal & Penelitian Teknik Informatika*, vol. 3, no. 1, pp. 272-278, 2018.

[9] P. B. I. Sukadiana Putra, N. P. S. M. Suryani and S. Aryani, "Analysis of Apriori Algorithm on Sales Transactions To Arrange Placement of Goods on Minimarket," *Internasional Journal of Engineering And Emerging Technology*, vol. 3, no. 1, pp. 13-17, 2018.

[10] R. Mujiastuti, P. Meilina and M. Anwar, "Implementasi Metode Economic Order Quantity (EOQ) Pada Sistem Informasi Produksi Kopi," *Jurnal Sistem Informasi, Teknologi Informasi dan Komputer*, vol. 8, no. 2, pp. 119-126, 2018.

[11] D. Nababan, "Sistem Pengontrolan Persediaan Barang Dengan Metode Economic Order Quantity (EOQ) menggunakan Algoritma Genetika," *Jurnal ISD*, vol. 2, no. 1, pp. 63-69, 2017.

[12] Z. A. Achmad and R. Ida, "Etnografi Virtual Sebagai Teknik Pengumpulan Data dan Metode Penelitian," *The Journal of Society dan Media*, vol. 2, no. 2, pp. 130-145, 2018.

[13] R. Mulyanti and A. Setiawan, "Market Basket Analysis Dengan ALgoritma Apriori Pada Ecommerce Toko Busana Muslin Trendy," *Junal Informatika*, vol. 8, no. 1, pp. 11-18, 2020.

[14] R. M. Hidayat, R. Muchlis and F. H. Badruzzaman, "Model EOQ untuk Permintaan Bergantung Waktu Dengan Mempertimbangkan Laju Kerusakan," *Jurnal Matematika UNISBA*, vol. 1, no. 1, 2016.