Conference on Electrical Engineering, Informatics, Industrial Technology, and Creative Media 2023

ANALYSIS OF VIRAL WARUNK SALES DATA USING THE APRIORI ALGORITHM METHOD TO DETERMINE CUSTOMER PURCHASING PATTERNS

Tri AristiSaputri#1, EkaPrasetiyo Budi*2

Information System, STMIK Dharma Wacana
Metro City - Lampung

[1]aristy@dharmawacana.ac.id

*T*Informatics Engineering, STMIK Dharma Wacana
Metro City - Lampung

[2]prasetiyoeka25@gmail.com

Abstract

This research aims to analyze Warunk Viral sales data using the Apriori algorithm method in order to determine consumer purchasing patterns. Warunk Viral is a small shop that sells various types of food and drink products. In this digital era, it is important for business owners like Warunk Viral to understand their consumer behavior in more depth in order to increase operational efficiency and profits.

The Apriori algorithm method is used in this research to identify products that are frequently purchased by consumers. By analyzing collected sales data, this research will reveal purchasing patterns that may not be immediately apparent, such as associations between certain products. The results of this analysis will provide valuable insight to Warunk Viral owners in managing stock, determining pricing strategies, and designing more effective promotions.

The research method used involves collecting Warunk Viral sales data over a certain period of time, processing the data, and applying the Apriori algorithm to generate relevant association rules. This data analysis can help Warunk Viral make smarter decisions in inventory management and improve the consumer shopping experience.

It is hoped that the results of this research can contribute to further understanding of consumer behavior at Warunk Viral and can also be applied in various other industries. In conclusion, this research proves that data analysis using the Apriori algorithm method can be an effective tool in understanding consumer purchasing patterns and helping companies to improve operational efficiency and increase their profits.

Keywords: Data Analysis, Viral Warunk, Apriori Algorithm Method, Consumer Purchasing Patterns, Purchasing Pattern Analysis

Corresponding Author:
SaadMekhilefPower Electronics and Renewable Energy Research Laboratory (PEAR-L), University of Malaya

BalaiCerap UTM, LengkokSuria, 81310 Skudai, Johor, Malaysia
E-mail:saad@um.edu.my

I.   INTRODUCTION

Sales is an integrated activity to develop strategic plans aimed at satisfying the needs and desires of buyers/consumers. The aim of sales is to bring in profits or profits from the products or services produced by the producer with good management and also expect large profits. Data Mining is a collection and filtering of data, using a very large collection of data with a series of processes aimed at obtaining important data information. So that sales data is useful, data collection is carried out so that we can find the types of products that are often purchased by consumers at the same time using a priori algorithm. The Apriori algorithm is a data collection algorithm that uses association rules to determine associative relationships from combinations of items. The target association rules are implemented using a mechanism that calculates the support and trust of the target relationship. Association rules are considered interesting if the support value is greater than the minimum support and the confidence value is also greater than the minimum confidence.

Warunk Viral is a culinary business that has been operating since the end of 2019. Warunk Viral offers 60 different food and drink menus. The cafe's operating hours are from 09.00 WIB to 22.00 WIB. Warunk viral is located on Jl. Soekarno Hatta, Mulyojati Village, West Metro District, Metro City, Lampung. Until now, Warunk Viral has served many transactions. All sales data has been stored in a database system via a management information system application. In this research the author tried to conduct an experiment using transaction data on Warunk Viral's food and drink menu. Warunk Viral has marketed its products using social media such as Facebook, Instagram, etc. However, this advertising is still not effective because Warunk Viral does not know which products are popular and which are not popular with consumers. Knowing all types of products that are frequently purchased can be the basis for making decisions aimed at determining which products are good to promote to customers. The research was carried out because Warunk Viral has problems, one of which is decreasing sales turnover. Therefore, Warunk Viral carries out promotions, but the promotions currently being implemented are less effective, because consumers are less interested in the advertising and information provided.

Based on the problems that have been explained, we need a method or technique that can provide information through a priori algorithm calculations. So it can be useful when carrying out promotions by using a priori algorithm data calculations to see which products are sold. The overall product is based on the results of the Support Score and Trust Score calculations and can help formulate sales strategies by creating promotional content for products that sell a lot and attract consumers.

1.1. SCOPE OF PROBLEM

So that cannot be separated from the background and problem formulation, the limitations of the problem in this research are:

1. The data used is sales transaction data in 1 August 2023 until 25 October 2023
2. Research uses an a priori algorithm method

1.2. RESEARCH PURPOSES

In accordance with the problem formulation and problem limitations, the goal to be achieved in this research is to apply data analysis to sales transactions at viral stalls using an a priori algorithm to determine consumer sales patterns.

II.   THEORETICAL BASIS

3.1 DATA MINING

Data Mining is an analysis step towards a process of discovering database knowledge or knowledge discovery in databases. Knowledge can also be in the form of relationships between

data or valid (previously unknown) data patterns. Data mining is a combination of a number of computer science disciplines which defines the process of discovering new patterns from a very large data set, including methods included in the intersection of AI (artificial intelligence), machine learning, statistics and database systems.

Data mining demonstrated with the aim of extracting knowledge from a data set so that it obtains a structure that is easy for humans to understand which includes database and data management, data processing, consideration of a model and inference, consideration of complexity, measurement of interest, post-processing of the structure found, online updating and visualization

3.2 APRIORI ALGORITHM

Apriori is an algorithm that is well known for searching frequent itemsets using association rule techniques. The a priori algorithm uses knowledge about previously known frequent itemsets to process further information. In the a priori algorithm to determine the candidates that may appear by paying attention to the minimum support [2].

The steps for the Apriori algorithm to obtain the rules desired by the user include:

(1) Scan the database to get a 1-itemset candidate, namely C1 (a set of items consisting of 1 item) and calculate the support value. Compare the support value with the minimum support that has been determined, if the value is greater than or equal to the minimum support, then the itemset is included in the large itemset, namely L1 (Large itemset with 1 itemset)
(2) Item sets that are not included in the large itemset are not included in the next iteration (pruning is carried out).
(3) The L1 set resulting from the first iteration will be used for the next iteration. In L1, a joining process is carried out on itself to form 2 candidate itemsets (C2). Compare again the support of C2 items with the minimum support, if it is not less than the minimum support, then the itemset is included in the large itemset L2. In the next iteration, the large itemset results in the previous iteration (Lk-1) will be joined by themselves to form a new candidate (Ck) and a new large itemset (Lk). After that, the pruning process is carried out on itemsets that are not included in Lk-.
(4) From all large itemsets that meet the minimum support (frequent itemsets), association rules and confidence values are formed. Rules whose confidence value is smaller than the minimum confidence are not included in the association rules used.

III.   RESEARCH METHODS

3.1 DATA COLLECTION

(1). INTERVIEWS

Researchers conducted interviews and collected data directly from viral warunk owners.

(2). OBSERVATIONS

Researchers made direct observations and carefully recorded the work system to collect

information and obtain the information needed, located at Warunk Viral Mulyojati,

West Metro District, Metro City, Lampung.

(3). LITERATURE STUDY

This previous research became one of the author's references in conducting research so

that the author could enrich the theory used in reviewing the conducted research. From

previous research, the author did not find research with the same title as the author's

research title. But the author raised several studies as references in enriching the study

material in the author's research. The following is previous research in the form of several journals related to research conducted by the author.

IV. RESULTS AND DISCUSSION

4.1 HIGH FREQUENCY PATTERN SEARCH ANALYSIS

This research data is sales transaction data for viral Warunk products which can be obtained through the cashier application (Fposystem). The research data used is sales from August 1 2023 to October 25 2023 with total sales of 86 days with a total of 3248 products sold from a total of 72 types of products owned by Warunk Viral.

4.2 CREATING DATA IN TABULAR FORM

NO	CODE	NAME	SUM	SUPPORT
1	A01	sosis	3	0.09
2	A02	Ayam bakar	105	3.23
3	A03	ayam goreng	4	0.12
4	A04	Ayam penyet	14	0.43
5	A05	Bebek bacem	15	0.46
6	A06	Bebek bakar	34	1.05
7	A07	Bebek goreng	23	0.71
8	A08	gurame bakar	126	3.88
9	A09	gurame asam manis	22	0.68
10	A10	Nila bakar	75	2.31
11	A11	Nila goreng	4	0.12
12	A12	Lele goreng	15	0.46
13	A13	Lele bakar	21	0.65
14	A14	Pindang baung	21	0.65
15	A15	Pindang patin	16	0.49
16	A16	Nasi goreng	187	5.76
17	A17	Ayam kampung	20	0.62
18	A18	Mie pangsit	96	2.96
19	A19	Seblak	53	1.63
…..	…..	……………..	…..	…..
…..	…..	……………..	…..	…..
71	B26	taro	14	0.43
72	B27	teh	137	4.22

Table I data is in tabular form

4.3 CREATING DATA IN BINARY FORM

TRANSACTION	sausage	grilled chicken pkt	butter fried chicken	…..	taro	tea
010823	0	1	0	…..	0	0
020823	0	0	0	…..	0	1
030823	0	0	0	…..	0	1
040823	0	0	0	…..	1	1
050823	0	0	0	…..	0	1
060823	0	1	0	…..	0	1
070823	0	1	0	…..	0	1
080823	0	1	0	…..	1	1
090823	0	1	0	…..	0	1
100823	0	1	0	…..	0	0
110823	0	1	0	…..	0	1
120823	0	0	0	…..	0	1
130823	0	1	0	…..	1	0
140823	0	1	0	…..	1	0
150823	0	0	0	…..	0	0
160823	0	0	0	…..	0	1
170823	0	1	0	…..	1	1
…..	…..	…..	…..	…..	1	1
…..	…..	…..	…..	…..	1	1
241023	0	0	0	……………	1	1
251023	0	0	0	……………	0	1

Table II transaction data in binary form

4.4 CALCULATING SUPPORT VALUE

The transaction data in binary form is then calculated for each support value and calculates the product combination in the viral warunk.

TRANSACTION	sausage	grilled chicken pkt	fried chicken	penyet chicken	bacem duck	…..	taro	tea
010823	0	1	0	0	0	…..	0	0
020823	0	0	0	0	0	…..	0	1
030823	0	0	0	0	0	…..	0	1
040823	0	0	0	0	0	…..	1	1
050823	0	0	0	0	0	…..	0	1
060823	0	1	0	1	0	…..	0	1
070823	0	1	0	0	0	…..	0	1

080823	0	1	0	0	0	.....	1	1
090823	0	1	0	0	0	.....	0	1
100823	0	1	0	0	0	.....	0	0
110823	0	1	0	0	0	.....	0	1
120823	0	0	0	0	0	.....	0	1
130823	0	1	0	1	0	.....	1	0
140823	0	1	0	0	0	.....	1	0
.....	.....	.....	.....	.....	.....	.....	.....	.....
.....	.....	.....	.....	.....	.....	.....	.....	.....
241023	0	0	0	0	0	.....	1	1
251023	0	0	0	0	0	.....	0	1
TOTAL	0	9	0	2	0	.....	5	12
SUPPORT	0	0.257142857	0	0.057142857	0	.....	0.14	0.34

Table III Determining the Support Value for Each Item

(1) LOOKING FOR HIGH FREQUENCY PATTERNS

In this stage, product combinations will be calculated

Support (A) = <u>Number of transactions containing A</u>

Total Transactions

Support(Degan)=72/86=0.837=83%

Support (B) = Number of transactions containing A

Total Transactions

Support(Nasi Goreng)=67/86=0.779=77.9%

(2) BUILDING ASSOCIATION RULES

The associations in this research determine 2 possible products that are frequently purchased and related. Retrieve associations based on the highest support value for drinks and the highest support value for food and test their association with each transaction.

4.5 CALCULATING CONFIDANCE VALUE

Sales is an integrated activity to develop strategic plans aimed at satisfying the needs and desires of buyers/consumers, in order to obtain sales that generate profit or profit. The definition of sales is a transaction activity carried out by 2 (two) or more parties using

Association A=>B

Confidance= P(B/A) = Number of transactions containing A and B

The transaction amount contains A

P(Degan& Nasi Goreng /Degan)=55/72=0.763=76.38%

P(With& Nasi Goreng / Nasi Goreng)=55/67=0.820=82%

From the results of the a priori algorithm calculations, the minimum support values were determined, namely 77.9 and 83%, with a confidence level of 76.38%, which is the trend for consumer purchasing products.

NO	CALCULATION TECHNIQUES	RESULTS
1	Calculations use an a priori algorithm	If consumers buy Nasi Goreng products, they will buy products with a confidence value of 76.38%
2		If consumers buy degan products, they will buy Nasi Goreng products with a confidence value of 82%

Table IV Calculation Results

The results of this research show some significant information related to consumer purchasing behavior at Warunk Viral. The following is a description of the results:

(1). Support for Degan Products (83%):

The support value of 83% indicates that the Degan product is one of the products that is quite popular at Warunk Viral. This means 83% of all purchase transactions involve Degan products. This shows that Degan products have high popularity among consumers.

(2). Support for Nasi Goreng Food (77.9%):

With a support value of 77.9%, Nasi Goreng food is also a product that is quite popular with Warunk Viral consumers. Approximately 77.9% of all purchase transactions involve Fried Rice. This indicates that Fried Rice is also a popular product at this restaurant.

(3). Association between Nasi Goreng Products and Degan:

The confidence value of 76.38% shows that when a consumer buys a Nasi Goreng product, there is a high probability (around 76.38%) that they will also buy a Degan product at the same time. This shows that there is a positive correlation between purchasing Nasi Goreng and Degan. In other words, these products are often purchased together by consumers.

(4). Association between Degan Products and Fried Rice:

A confidence value of 82% shows that when a consumer buys a Degan product, there is a high probability (around 82%) that they will also buy a Nasi Goreng product at the same time. this indicates a reciprocal relationship between these two products, where purchasing one product increases the likelihood of purchasing the other.

Degan and Fried Rice products have high popularity in Warunk Viral. In addition, there is a strong purchasing pattern between these two products, where the purchase of one product is often followed by the purchase of the other product. This information can be useful in designing marketing strategies, preparing menus, or managing product stock at Warunk Viral to increase sales and consumer satisfaction.

Because these two products have a high level of support, it is recommended to increase their promotion and marketing, by creating promotional packages that combine these two products or providing incentives to customers to buy both together. This will drive sales and increase revenue.

V. CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusion:
1 This research has succeeded in identifying significant purchasing patterns on Warunk Viral. Degan and Nasi Goreng products stand out as popular products among consumers, with high levels of support.
2 There is a strong association between purchasing Degan and Nasi Goreng. Consumers who buy one product tend to also buy the other product, indicating a positive relationship between these two products.
3 The results of this research provide a great opportunity for Warunk Viral to increase sales by optimizing their marketing and menu strategy. Combining Degan and Nasi Goreng products in promotions or creating special menus can increase the appeal for customers.

5.2 Suggestion:
1 Warunk Viral owners should consider increasing promotion of Degan and Nasi Goreng products. This can be done through various strategies, including discounts, promo packages, or special promotions for this product.
2 Continue to carry out regular data analysis to monitor sales trends and ensure the strategies implemented are still effective. If there is a change in purchasing patterns, act immediately with an appropriate strategy.

REFERENCES

[1] Bima Arif Saputra, B. S. (2023). APLIKASI SISTEM PAKAR DIAGNOSA PENYAKIT GIGIBERBASIS WEBSITE MENGGUNAKAN METODE TEOREMA BAYES. *BULLETIN OF NETWORK ENGINEER AND*, 33-41.

[2] Didi Susianto Yuli syafitri, G. Y. (2022). *Sistem Informasi Manajemen.* Adab.

[3] Fadila Shely Amalia, S. D. (2021). Analisis Data Penjualan Handphone dan Elektronik menggunakan Algoritma Apriori (Studi Kasus :CV Rey Gasendra). *Telefortech*, 1-6.

[4] Lia Aprita, A. P. (2023). Penerapan Metode Data Mining terhadap Data Transaksi Penjualan Menggunakan Algoritma Apriori pada Toko Metro Akustik. *Jurnal Teknologi Informatika dan Komputer*, 274-283.

[5] Muhammad Briliantino, A. P. (2023). Penerapan Algoritma Apriori pada Analisis Data Transaksi penjualan UMKM Banyu Burgerbar. *Jurnal Teknologi Informatika dan Komputer*, 61-71.

[6] Ridwan Yusuf, T. A. (2021). PENERAPAN NATURAL LANGUAGE PROCESSING BERBASIS VIRTUAL ASSISTANT PADA BAGIAN ADMINISTRASI AKADEMIK STMIK DHARMA WACANA. *International Research on Big-Data and Computer Technology: I-Robot*, 33-47.

[7] Ririn Restu Aria, S. S. (2019). Analisa Data Penjualan SaRa Collection menggunakan metode Apriori. *Jurnal Teknik Komputer AMIK BSI*.

[8] Saefudin, S. (2019). Penerapan Data Mining Dengan Metode Algoritma Apriori untuk Menentukan Pola Pembeian Ikan. *Jurnal Sistem Informasi*, 110-114.

[9] Thaariq Nasrah, K. N. (2021). Penerapan Algoritma Apriori Pada Penjualan Kopi Arabica. *Semnastek UISU*.

[10] Zaenal Abidin, A. K. (2022). Penerapan Algoritma Apriori Pada Penjualan Suku Cadang Kendaraan Roda Dua (Studi Kasus : Toko Prima Motor Sidomulyo). *Jurnal Teknoinfo*, 225-232.