2021 IEEE 5th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE) | 978-1-6654-0196-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICITISEE53823.2021.9655794

ISBN 978-1-6654-0196-8
IEEE Catalog Number CFP21G48-ART

ICITISEE 2021

2021 IEEE 5th INTERNATIONAL CONFERENCE ON INFORMATION TECNOLOGY, INFORMATION SYSTEM AND ELECTRICAL ENGINEERING (ICITISEE)

PROGRAM BOOK

November 24th - 25th 2021

"Applying Data Science and Artificial Intelligence Technologies for Global Challenges During Pandemic Era"

KEYNOTE :
Professor Mohammed Lazaar, Ph.D. (Mohammed V University, Morocco)
Dr.Eng. Sunu Wibirama, M.Eng. (Universitas Gadjah Mada, Indonesia)
Dr Eng. Imam Tahyudin, M.M.(Universitas Amikom Purwokerto, Indonesia)

Organized by :

UNIVERSITAS AMIKOM PURWOKERTO

Technical co-sponsored :

IEEE INDONESIA SECTION

Supported by :

The Indonesia Section Computer Society Chapter

Sponsored by :

Bank Muamalat
BANK BRI
mandiri
GMEDIA
Multimatics

PROCEEDINGS

2021 IEEE 5th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)

24-25 November 2021

Universitas Amikom Purwokerto, Purwokerto, Indonesia

2021 IEEE 5th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE) | 978-1-6654-0196-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICITISEE53823.2021.9655802

COPYRIGHTS

2021 IEEE 5th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)

IEEE Catalog Number : CFP21G48-ART
ISBN : 978-1-6654-0196-8

Additional copies of this publication is available from

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA

+1 845 758 0400
+1 845 758 2633 (FAX)
Email: curran@proceedings.com

2021 IEEE 5th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE) | 978-1-6654-0196-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICITISEE53823.2021.9655912

ICITISEE

Table of Content

2021 IEEE 5th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE) | 978-1-6654-0196-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICITISEE53823.2021.9655892

Analysis of Customer Purchase Behavior using Association Rules in e-Shop

1st Saucha Diwandari
Medical Informatics
Universitas Teknologi Yogyakarta
Yogyakarta, Indonesia
saucha.diwandari@staff.uty.ac.id

2nd Umar Zaky
Information System
line 3: *Universitas Teknologi Yogyakarta*
Yogyakarta, Indoenesia
umar.zaky@staff.uty.ac.id

Abstract— Modern forms of trade and sales support systems are meant to give organizations flexibility when it comes to evaluating sales. It is essential to identify customer buying behavior to predict customer intent. Web usage mining is a method that can be used in this regard. The research aimed to find customer segmentation to help businesspeople identify products that customers are interested in, and the right strategy to increase the chances of achieving competitive advantage using the Apriori and FP algorithms. The research results showed that the Apriori algorithm and FP-Growth helped conduct evaluations, especially to identify which products are of interest to customers. It found that out of 217 products, only 3 had the certainty of being purchased by customers. This information can help businesses to focus on selling specific products.

Keywords—association, web usage mining, web log data, user behavior, data mining

I. INTRODUCTION

Businesses may use online activities to better understand their customers' behavior on e-commerce platforms. Consumer-related knowledge can help a company anticipate buyer needs and provide personalized services according to buyer preferences [1]. Data mining techniques, particularly online use mining, make it simple to collect and analyze precise information about client activity in an electronic environment [2]. The information obtained will be put to use in the e-shop to improve customer service, attract new visitors, increase customer satisfaction to increase competitive advantage. In web usage mining, several techniques can be used to find user patterns, including classification [3][4], clustering, and association rule [2]. Association rule is a popular method used to explore information on the Internet. Discovering association rules, also known as market basket analysis, is a technique for finding association rules between a combination of items [5]. Association rule refers to pattern discovery and clickstream analysis with other variables collected or generated during customer interactions with the same or different websites. The patterns and models created are usually presented in a set of pages, objects, or anything that has the highest frequency of access [6]–[9]. In this paper, we propose discovering association rules with various user session features related to product purchases. We identify customer buying behavior in Poland to predict customer intent using two algorithms, namely Apriori and FP-Growth. The research aims to find customer segmentation to help businesspeople identify products of interest to customers and find the right strategy to increase opportunities to achieve competitive advantage.

II. RELATED WORK

Several researchers have carried out studies using association rule mining methods [2],[8], [10]–[13]. According to G. Suchacka et al. [2], using the association rule method to find the probability of purchase in a user session depends on the product category viewed and the session id on the online bookstore. The results of this study indicate that users who view printed book products through an online bookstore for 10 to 25 minutes and have opened a website page between 30 to 75 pages have a probability of buying a book more than 92%. X. Guo et al. [10] utilized association rules to extract connectivity and correlation of passenger flows between various service lines in the city rail transport network. More important association rules (2-frequent itemset and 3-frequent itemset) provide a better understanding of the transfer behavior of city train passengers. In contrast, research with this paper [11] compared several methods in association rule mining to study the behavior of online shoppers and predict whether they will buy a product or not. In websites, the Apriori algorithm, which is from frequent pattern mining, can also provide recommendations regarding exciting links to be opened by customers in e-commerce [13]. Another study [14] used association rules to predict human activity using temporal sequence patterns in accessing online videos. Various association algorithms are used, such as Rapid Association Rule Mining (RARM), Equivalence class clustering, Eclat, diffset Eclat (dEclat), and Frequent Pattern-growth (FP-growth). Association rule mining is used to predict and suggest web pages that are likely to be visited by web users; the data includes 2,370 web pages accessed by 77 different users [12]. It indicates that the association algorithm can be used to predict web pages. Those studies show that association rule mining can be used for various purposes. Thus, in this study, the Apriori and FP-Growth algorithms performance analysis on the E-Shop web-log data will be compared to find out which algorithm is the best.

III. RESEARCH METHODOLOGY

Web use mining (WUM) is a technique for analyzing weblog data patterns to gain a better understanding of web users. Dataset collection, pre-processing, pattern identification, and pattern analysis are the four steps of the WUM approach [15]. We established the proposed conceptual framework for analysis, depicted in Figure 1. There are six stages to the structure. The types of datasets are defined in the first stage. Stage 2 divides the input data in those datasets into dimensions. Stage 3 depicts the approaches used to create features from the incoming data and how customer behavior is modeled for predictive analytics. The prediction model is the fourth stage. Stage 5 shows tasks that enable the application from stage 6.

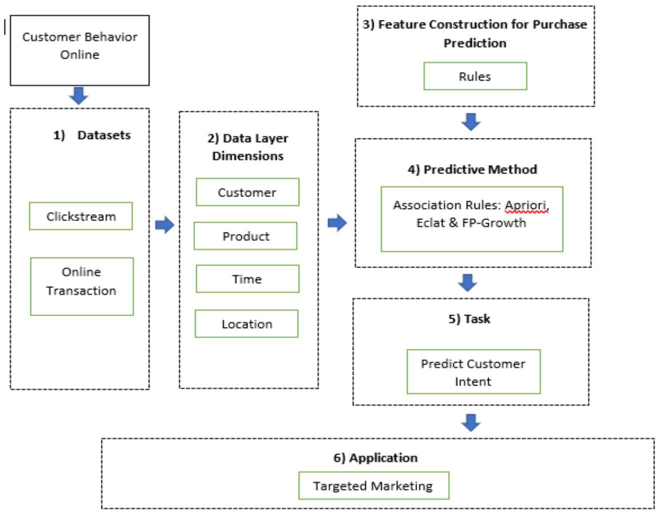

Fig. 1. A framework of analysis for customer purchase behavior

A. Collecting Datasets

- Datasets of previous online sessions and shopping logs are used to track customer behavior in e-commerce.

1) Transaction. Purchases executed by customers within the E-commerce. Data layer dimensions involved, namely customer, product, time, channel, and location
2) Customer: reveals the profile of every consumer and enables their segmentation. Input data features, namely demographics, clicks, session variables. A session is a series of pages viewed by the users during a specific visit.
3) Product: refers to the internet product's raw properties. Customers' purchasing information aids in recognition of preferences based on product features. The input data's value, description, and availability status are all characteristics.
4) Time: timestamps of customer transactions. Input data feature, namely timestamp and season
5) Channel: their characteristics define touchpoints between customers and an e-commerce platform. You can determine the value of a customer's purchases across several channels by knowing the customer's channel. Customer device and visit source are two types of input data
6) Location: consumer location data can assist in finding patterns based on customer geographical positioning. Input data features namely neighborhood and city

B. Frequent Itemset Mining

The purpose of frequent itemset mining is to identify recurring groupings of items in a database containing these transactions. Several different methods have been employed to mine frequent item-sets. This paper demonstrates how Apriori and FP-Growth can find comparable consumer groups and applications that require customer segmentation. All two methods must compute statistics for item-sets included in the final collection of often occurring item-sets. Support is a statistic that all four of these methods have in common. The overall count of how many database baskets support a candidate's frequent itemset is the candidate support. Support is sometimes expressed as a parameter that indicates how often an item appears in the dataset. The number of transactions in the data that contain item X is defined as the support of item X about a transaction T from a mathematical stand point[5]. The parameter of support for a candidate itemset is calculated using the formula below:

$$\sigma\ (X) = |\{t\ \in T; X \subset t\}| \tag{1}$$

In which $\sigma\ (X)$ is the Support Count. Support is a parameter that describes how frequently an item appears in a transaction. If the total number of transactions is 'N,' then X support is defined as

$$Support\ (X \rightarrow Y) = \frac{\sigma(X \cup Y)}{N} \tag{2}$$

The apriori method has been around for almost as long as the notion of frequent item sets has been, and it is quite popular. Jeff's definition of apriori [8] is presented as Algorithm 1.

TABLE I. APRIORI FREQUENT ITEMSET ALGORITHM

Algorithm 1 Apriori Frequent Itemset Algorithm
INPUT: A file D consisting of baskets of items, a support threshold σ.
OUTPUT: A list of item-sets $\mathcal{F}(D, \sigma)$.
METHOD:
C1 ← $\{\{i\} \mid i\ \in \mathcal{J}\}$
k ← 1
while $C_k \neq \{\}$ **do**
# Compute the supports of all candidate item-sets
for all transactions $\{tid, I\} \in D$ **do**
for all candidate item-sets $X \in C_k$ **do**
If $X\ \subseteq I$ then
$X.support++$
#Extract all frequent item-sets
$F_k = \{X \mid X.support\ > \sigma\}$
# Generate new candidate itemset
for all $X, Y\ \in Fi, X[i] = Y[i]\ for\ 1 \leq i \leq k-1$ and $X[k] < Y[k]$ **do**
$I = X\ \cup \{Y[k]\}$
If $\forall J \subset I, \lvert J \rvert = k{:}\, J \in F_k$ **then**
$C_{k+1} \leftarrow C_{k+1} \cup I$
$k++$

Big-O analysis was not used to express the FP-Growth method. Jeff's definition of FP-Growth[8], is presented as Algorithm 2.

TABLE II. FP-GROWTH ALGORITHM

Algorithm 3 FP-Growth Frequent Itemset Algorithm
INPUT: A file consisting of baskets of items, a support threshold σ, and an item prefix I, such that $I \subseteq \mathcal{J}$
OUTPUT: A list of item-sets $\mathcal{F}[i](D, \sigma)$ for the specified prefix
$\mathcal{F}[i] \leftarrow \{\}$
for all $i\ \in \mathcal{J}$ occurring in D **do**
$\mathcal{F}[I] := \mathcal{F}[I] \cup \{I \cup \{i\}\}$

```
# Create D_i
  D_i ← {}
  H ← {}
  for all j ∈ J occurring in D do
     if support (I ∪ {i, j}) ≥ σ then
        H ← H ∪ {j}
  for all (tid, X) ∈ D with I ∈ X do
        D_i ← D_i ∪ {j, C}
  #Depth-first recursion
     Compute F[I ∪ i](D_i)
       F[I] := F[I] ∪ F[I ∪ i]
```

IV. RESULT & ANALYSIS

The data collected included 210,814 logs from web log data in https://archive.ics.uci.edu/. The study relied on a set of log files from an online store that sold women's apparel. The 210,814 entries in the dataset were related to spring and summer clothing in the areas of skirts and dresses, trousers, blouses, and special offers. The website was available in various languages, including English, German, and Czech, allowing potential consumers from beyond Poland to access it. The firm agreed to ship internationally, making it easier to fulfill requests from different parts of the world. Visitors' IP addresses were collected from site logs, and each address in the database was allocated to a specific nation based on geolocation. Due to several challenges, determining the user's nation of residence was not entirely achievable. In certain circumstances, domain names terminated in the letters.net,.com,.biz, or.org, and they were also recorded as text rather than numbers. The geolocation system would sometimes only show a continent, such as Europe, rather than a single nation. Table 3 presents the weblog file format used.

TABLE III. EXAMPLE OF THE WEB LOG DATA

Field	Meaning
Year	2008
Month	4
Day	1
Order	1
Country	29
Session ID	1
Main Category	1
Clothing Model	A13
Colour	1
Location	5
Model Photography	1
Cost Cost 2	28 2
Page	1

With a description of the data as follows:

1. Year
2. Month → April (4) to August (8)
3. Day → day number of the mount
4. Order → sequence of clicks during one session
5. Country → IP address with the following categories as a variable denoting the country of origin: 1) Australia, (2) Austria, (3) Belgium, (4) British Virgin Island, (5) Cayman Islands, (6) Christmas Island ect
6. Session ID → varible indicating session id (short record)
7. Page 1 (Main Category) → concerns the main product category : (1) trousers, (2) skirts, (3) blouses, (4) sale
8. Page 2 (Clothing Model) → contains information about the code for each product (217 product)
9. Colour → (1) beige, (2) black, (3) blue, (4) brown, (5) burgundy, (6) gray, (7) green, (8) navy blue, (9) of many colors, (10) olive, (11) pink, (12) red, (13) violet dan (14) white
10. Location → The screen has been split into six pieces based on the position of the photo on the page: (1) top left, (2) top in the middle, (3) top right, (4) bottom left, (5) bottom in the middle, (6) bottom right
11. Model photography → variable with two categories: (1) en face, (2) profile
12. Cost → cost in US dollars
13. Cost 2 → variable informing whether the cost of a particular product is higher than
    the average price for the entire product category: (1) yes, (2) no
14. Page → the e-page shop's number (from 1 to 5)

A. Modelling

In this study, two algorithms are used to form rules on sales transaction data. The a priori method aims to identify an association rule that meets the minimal support (supporting value), representing the combination of each item in the database, and the confidence (certainty value), which means the strong link between objects in the association rule. The FP-Growth method may directly extract frequent itemset from FP-Tree using FP-Tree

1) *Formation of 1-itemset*: In this study, transaction data was used. Because the minimum number supplied is 100, the method of creating C1, also known as 1-itemset, involves calculating the frequency of occurrence of item-sets or frequent item-sets in transaction data and setting the minimum value. [9][16]. Thus, out of 217 products, only 15 products are used from total transactions collected which were 903 in Poland. The types of products that fulfill the minimum value of a frequent itemset or the required occurrence frequency are as follows:

TABLE IV. FREQUENT ITEMSET

Itemset	Frequent
A1	132
A2	149
A4	135
A5	116
A11	133
A15	141
A17	100
B4	180
B10	139
B13	101
C5	105
P1	183
P6	110

P15	107
P16	102

2) *Combination of 2-Itemset:* Constructing C2, identifying the frequency of occurrence of a combination of two items in transaction data and estimating the value of its support, also known as a 2-itemset (equation 1). With 0.04 as an example, the following combination fulfills the stated minimum support.

TABLE V. MINIMUM SUPPORT AND CONFIDENCE 2-ITEMSET

Itemset	Frequent	Support
A17, A5	36	0.030
A17, A1	24	0.026
A17, A11	25	0.027
A17, A4	20	0.022
A17, A15	29	0.032
A17, A2	23	0.025
A17, B4	22	0.024
A5, A1	37	0.040
A5, A11	29	0.032
A5, A4	25	0.027
A5, A15	21	0.023
A5, A2	49	0.054
A5, B4	19	0.021
A1, A11	43	0.047
A1, A4	36	0.039
A1, A15	33	0.036
A1, A2	55	0.060
A1, B4	30	0.033
A11, A4	28	0.031
A11, A15	30	0.033
A11, A2	38	0.042
A11, B4	28	0.031
A4, A15	37	0.040
A4, A2	28	0.031
A4, B4	29	0.032
A15, A2	32	0.032
A15, B4	27	0.029
A2, B4	24	0.026

3) *Combination of 3-Itemset* : C3, also known as the 3-itemset, assesses the value of support by determining the frequency of occurrence of a combination of three items in transaction data.
4) Formation of Association Rules: Forms association rules govern how itemsets on transactions, such as confidence, are combined. The equation below can be used to solve it (2)
5) Calculation of the confidence value for 2-itemset association rules, then the minimal confidence, and finally the final association calculation.
6) Lift Ratio: The lift ratio is a measure to find out the strength of the association rule has formed. The lift ratio value is used to determine whether the association rule is valid or not.*Testing using RapidMiner*

In this stage, testing is carried out, aiming at:

a) FP-Growth and the influence of little support on the number of frequent item-sets created in the a priori method

b) Analyzing the effect of minimum confidence on the number of rules generated based on frequent itemset

c) To find out the strength of the association rules that have been formed through the lift ratio

The results of the analysis of the effect of minimum support and minimum confidence on frequent itemset can be seen in Table VI

TABLE VI. FREQUENT ITEMSET APRIORI ALGORITHM

Min Confidence	Min Support		
	30%	60%	85%
50%	-	-	-
70%	-	-	-
90%	-	-	30

TABLE VII. FREQUENT ITEMSET FP-GROWTH ALGORITHM

Min Confidence	Min Support		
	30%	60%	85%
50%	1023	731	137
70%	1023	731	137
90%	1023	731	137

Tables VI and VII show that the numbers of frequent item-sets produced by the two algorithms have differences. In the Apriori algorithm, the frequent itemset results can only be generated if the min support is 85% and the min confidence 90% is 30. In the FP-Growth algorithm, the frequent itemset results can be formed in various min support and min confidence. It is presented in Table VII that the most frequent itemset is generated from the smallest minimum support, which is 30%, while the least frequent itemset is from the minimum support of 85%. It is because the minimum support value is a reference in calculating an item set. Thus, the higher the minimum support used, the less frequent itemset will be generated.

TABLE VIII. NUMBER OF RULES FROM APRIORI ALGORITHM

Min Confidence	Min Support		
	30%	60%	85%
50%	-	-	-
70%	-	-	-
90%	-	-	10

TABLE IX. NUMBER OF RULES FROM FP-GROWTH ALGORITHM

Min Confidence	Min Support		
	30%	60%	85%
50%	5950	4442	986
70%	5592	4306	986

Min Confidence	Min Support		
	30%	60%	85%
90%	940	805	283

TABLE X. COMPARISON OF FP-GROWTH AND APRIORI RULE RESULT (MIN SUPPORT 85% AND MIN CONFIDENCE 90%)

Algorithm	Rule	Confidence	Lift Ratio
FP-GROWTH	**IF** Buy A17, A11, A2 **THEN** Buy A5, A1	0.911	1.082
	IF Buy A17, A4, A2 **THEN** Buy A5, A1	0.909	1.079
	IF Buy A1, A15, A2 **THEN** Buy A17, A5	0.910	1.068
APRIORI	**IF** Buy A2, A11 **THEN** Buy A1	1	1.03
	IF Buy A2, A4, A15 **THEN** Buy A1	1	1.02
	IF Buy A4, A5, A11 **THEN** Buy A1	1	1.01

Table VIII and Table IX present the results of comparing the number of rules generated in the two algorithms. In the Apriori algorithm, ten rules are developed at a minimum support of 85% and minimum confidence of 90%. Unlike the case with the FP-Growth algorithm, for every increase in the value of the minimum support and minimum confidence tested, the rules formed are decreasing. In this test, the high minimum support and minimum confidence used will generate fewer rules. Furthermore, testing is carried out on the lift ratio results for each rule formed. In this test, we only take the rule formed from minimum support of 85% & a minimum of 90% confidence in Table X.

The table presents the three rules with a lift ratio greater than 1 (lift ratio > 1), which means the resulting rule is valid. In the two algorithms above, there are differences in each of the resulting rules. It is determined by the difference between the concepts of the Apriori method and FP-Growth in determining the frequent itemset. In addition, there are also differences in the confidence results formed where the Apriori algorithm can reach 1. In contrast, the FP-Growth algorithm has the highest confidence value of 0.911. A confidence value that is higher or equal to 1 indicates the level of certainty of the occurrence of several products being purchased concurrently where one product is purchased. Overall, from the rules formed between yahoo Apriori and FP-Growth, it is known that if A11, A2 products are purchased, A1 products will also be purchased concurrently.

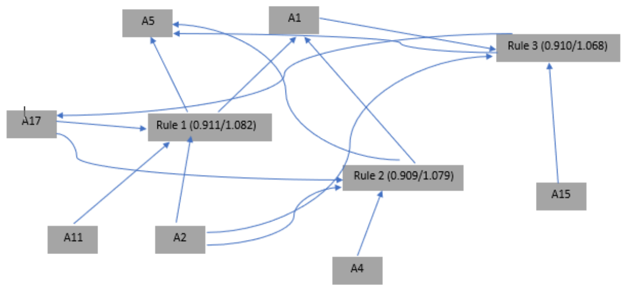

Fig. 2. A Graph Experiment Result of FP-Growth

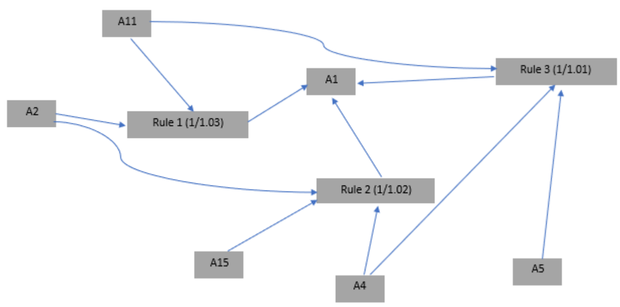

Fig. 3. A Graph Experiment Result of Apriori

V. CONCLUSION

The pattern of customer behavior in an e-Shop obtained from weblog data can be identified through association rules. According to the research results using the Apriori and FP algorithms, three rules have a high level of rule association for each algorithm. The consumer will purchase certain products concurrently. For example, if products A11 and A2 are purchased, product A1 will be purchased as well. Knowing the product information purchased can help business owners increase sales on certain products in high demand among customers. It would be beneficial to continue the experiment in the future by comparing consumer groups.

REFERENCES

[1] M. Ali *et al.*, "Customer Opinion Mining by Comments Classification using Machine Learning," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 5, pp. 385–393, 2021, doi: 10.14569/IJACSA.2021.0120547.

[2] G. Suchacka and G. Chodak, "Using association rules to assess purchase probability in online stores," *Information Systems and e-Business Management*, vol. 15, no. 3, pp. 751–780, 2017, doi: 10.1007/s10257-016-0329-4.

[3] S. Diwandari, A. E. Permanasari, and I. Hidayah, "Research methodology for analysis of E-commerce user activity based on user interest using web usage mining," *Journal of ICT Research and Applications*, vol. 12, no. 1, pp. 54–69, 2018, doi: 10.5614/itbj.ict.res.appl.2018.12.1.4.

[4] S. Diwandari and A. T. Hidayat, "Comparison of Classification Performance Based on Dynamic Mining of User Interest Navigation Pattern in e-Commerce Websites," *Journal of Physics: Conference Series*, vol. 1844, no. 1, p. 012025, Mar. 2021, doi: 10.1088/1742-6596/1844/1/012025.

[5] D. Mohapatra, J. Tripathy, K. K. Mohanty, and D. S. K. Nayak, "Interpretation of Optimized Hyper Parameters in Associative Rule Learning using Eclat and Apriori," 2021, pp. 879–882. doi: 10.1109/ICCMC51019.2021.9418049.

[6] M. Łapczyński and S. Białowąs, "Discovering Patterns of Users' Behavior in an E-shop - Comparison of Consumer Buying Behaviours in Poland and Other European Countries," *Studia Ekonomiczne*, vol. 151, pp. 144–153, 2013.

[7] B. Zheng and B. Liu, "A scalable purchase intention prediction system using extreme gradient boosting machines with browsing content entropy," 2018, vol. 2018-January, pp. 1–4. doi: 10.1109/ICCE.2018.8326351.

[8] J. Heaton, "Comparing dataset characteristics that favor the Apriori, Eclat or FP-Growth frequent itemset mining algorithms," 2016, vol. 2016-July. doi: 10.1109/SECON.2016.7506659.

[9] Islamiyah, P. L. Ginting, N. Dengen, and M. Taruk, "Comparison of Priori and FP-Growth Algorithms in Determining Association Rules," 2019, pp. 320–323. doi: 10.1109/ICEEIE47180.2019.8981438.

[10] X. Guo, D. Z. W. Wang, J. Wu, H. Sun, and L. Zhou, "Mining commuting behavior of urban rail transit network by using association rules," *Physica A: Statistical Mechanics and its Applications*, vol. 559, 2020, doi: 10.1016/j.physa.2020.125094.

[11] R. Obiedat, "A comparative study of different data mining algorithms with different oversampling techniques in predicting online shopper behavior," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 9, no. 3, pp. 3575–3583, 2020, doi: 10.30534/ijatcse/2020/164932020.

[12] Y. Raju, D. Suresh Babu, and K. Anuradha, "A web usage association learning methods based on mining techniques for web personalization," *Journal of Theoretical and Applied Information Technology*, vol. 96, no. 10, pp. 3116–3126, 2018.

[13] L. Zheng, "Research on e-commerce potential client mining applied to apriori association rule algorithm," 2020, pp. 667–670. doi: 10.1109/ICITBS49701.2020.00146.

[14] D. Manju and V. Radha, *Mining temporal sequence patterns using association rule mining algorithms for prediction of human activity from surveillance videos*, vol. 1200 AISC. 2021, p. 483. doi: 10.1007/978-3-030-51859-2_43.

[15] D. Cirqueira, M. Hofer, D. Nedbal, M. Helfert, and M. Bezbradica, *Customer purchase behavior prediction in e-commerce: A conceptual framework and research agenda*, vol. 11948 LNAI. 2020, p. 136. doi: 10.1007/978-3-030-48861-1_8.

[16] S. Vijayarani and S. Sharmila, "Comparative analysis of association rule mining algorithms," in *2016 International Conference on Inventive Computation Technologies (ICICT)*, Aug. 2016, vol. 3, pp. 1–6. doi: 10.1109/INVENTIVE.2016.7830203.