MATHEMATICAL MODELING AND COMPUTING, Vol. 10, No. 2, pp. 547–556 (2023)

Revolutionizing supermarket services with hierarchical association rule mining

Meftah M.[1], Ounacer S.[1], Ardchir S.[2], El Ghazouani M.[1], Azzouazi M.[1]

[1]*Laboratory of Information Technology and Modeling, Hassan II University, Faculty of Sciences Ben M'Sik, Casablanca, Morocco*
[2]*Hassan II University, National School of Commerce and Management, Casablanca, Morocco*

(Received 19 February 2023; Accepted 4 April 2023)

The use of association rule mining techniques has become a focal point for many researchers seeking a better understanding of consumer behavior. By analyzing the relationships between products and their placement in aisles, valuable insights can be gained into the factors that influence product preservation in large-scale distribution environments. This approach has the potential to inform better decision-making processes and optimize product preservation outcomes, despite some limitations in the quality of the data available. Additionally, a hybridization approach was adopted by incorporating transactions from clients participating in a loyalty program to encourage large-scale distributions to gain a better understanding of customer behavior and improve their purchasing strategies. The goal of this research is to promote consistency between the real-world and virtual representations of customer behavior, ultimately leading to improved purchasing outcomes for large-scale distributions.

Keywords: *product preservation; hybridation approach; loyalty program; association rules; apriori; large-scale distribution.*

2010 MSC: 90B60, 68P20, 91B42, 62P30 **DOI:** 10.23939/mmc2023.02.547

1. Introduction

The examination of customer behavior in large-scale distribution environments is a significant area of concern for companies aiming to gain a more profound knowledge of their clientele and optimize their processes. Association rule mining is a valuable technique for discovering valuable insights into customer behavior, which involves identifying the correlations between items in a market basket. Association rule mining is a data mining technique that uses advanced algorithms, such as the Apriori algorithm, to identify patterns and relationships in large datasets. By understanding the items that customers tend to purchase together, retailers can inform their marketing strategies and improve their inventory management. For example, they can use this information to decide which products to sell or promote together, or to design loyalty programs and sales promotions. However, there are several important factors to consider when utilizing association rule mining in large-scale distribution settings. One of these factors is the preservation conditions of the products being analyzed. These conditions can impact the quality of the products and affect customer behavior, and it is important to take them into account when identifying relationships between items in a market basket. Another important factor to consider is the information from loyalty program participants. This information can provide valuable insights into customer behavior and can be incorporated into decision-making processes for large-scale distributors. For example, it can help retailers to understand the preferences and habits of their most loyal customers and to design marketing strategies that are tailored to these customers. In conclusion, association rule mining is a powerful tool for improving our understanding of customer behavior in large-scale distribution settings. By considering factors such as the preservation conditions of products and incorporating information from loyalty program participants, we can gain deeper insights into customer behavior and make more informed decisions. The aim of this study is to explore how association rule mining techniques can enhance our understanding of customer behavior in large-scale distribution settings, and to investigate the impact of accounting for product preservation conditions and integrating information from loyalty program participants. The following research question is addressed: How can association rule mining techniques be applied to analyze customer behavior, and what is the effect of considering preservation conditions and loyalty program data?

The field of association rule mining has evolved greatly since the introduction of the first algorithm by R. Agrawal et al. in 1993 [1]. Over the years, researchers have developed and compared various algorithms, such as Apriori, AprioriTid, AprioriHybrid, RAC, MultipassApriori, Multipass-DHP, Fuzzy Association Rules, CBAR, TopkRules, HUCI-Miner, and association rules with utility. These algorithms have been used in various applications, such as text mining, data cleaning, and market basket analysis. The most recent development in this field is an association-based method for cleaning up outliers developed by H. Kuang et al. in 2021 [2].

To answer the research question, a comprehensive literature review will be conducted to gather information on various association rule mining techniques and their applications in large-scale distribution settings such as supermarkets or retail chains. The challenges and limitations of these techniques will also be investigated, including considerations for product preservation and the analysis of purchasing behavior for loyalty program participants. Based on these findings, a proposed solution will take into account the preservation conditions of products and the incorporation of loyalty program participants, providing a comprehensive approach to association rule mining in large-scale distribution settings.

The paper is structured as follows. Section 1 provides an introduction to the problem and research question that will be addressed in this study. In Section 2, classical techniques for association rule mining and their limitations are reviewed. Section 3 presents an approach that considers the preservation conditions of products and leverages data from customers engaged in loyalty programs to identify item associations. This section also explains the importance of preserving the quality of products and the role of loyalty programs in this context. In Section 4, the dataset used in the study, the algorithm used, and the results of the analysis are described, and the findings are discussed. Section 5 provides a conclusion and highlights the effectiveness of the approach using real-world data.

2. Related work

As the 20th century drew to a close, researchers began exploring the potential of big data and data mining. One significant breakthrough was the introduction of the first algorithm for association rule mining by Agrawal R. et al. in 1993 [1].

Subsequently, R. Agrawal and R. Srikant developed two new mining association rules algorithms, Apriori and AprioriTid, which outperformed previous algorithms such as AIS and SETM. They also suggested a hybrid algorithm, AprioriHybrid, which became the algorithm of choice for mining association rules on very important databases [3].

In 1998, B. Liu et al. integrated classification and association rules in an algorithm that generated all class association rules (RAC) and developed an accurate classifier. This solution addressed several problems that existed in current classification systems [4].

D. Holt and M. Chung developed two new algorithms, MultipassApriori and Multipass-DHP, in 1999 to detect hidden associations between words in textual databases. These algorithms were compared with existing algorithms such as Direct Hashing and Pruning, and Apriori, and were proven effective in large textual databases [5].

G. Chen and Q. Wei introduced fuzzy rules of association with linguistic barriers to express meaningful knowledge in a more natural and abstract manner in 2002 [6].

X. Wu et al. developed an effective method for extracting positive and negative association rules using a pruning strategy and a measure of interest in 2004 [7].

Yang et al. used Hadoop's MapReduce model for parallelizing Apriori to process large datasets with high numbers of nodes on the Hadoop platform in 2010 [8].

In 2012, Fournier-Viger et al. suggested an algorithm known as TopkRules that solved the problem of generating too many rules by mining association rules [9].

J. Sahoo et al. proposed a compressed representation for association rules with minimum antecedent and maximum consequence, generated using high utility closed itemsets (HUCI) and their generators in 2015. They suggested algorithms for generating non-redundant association rules and utility-based methods for rebuilding all association rules, which were implemented using synthetic and real datasets [10].

J. Wu et al. suggested a new approach that used a different combination of high-frequency and low-frequency objects, combining utility with frequency to derive different association rules to properly process low frequency elements in a transaction database in 2018 [11].

Y. Ünvan used association rules to carry out a market basket analysis in 2020 and observed that using the Apriori and FP Growth algorithms to generate rules allowed for better product placement in supermarkets and resulted in increased sales and revenue [12].

In 2021, H. Kuang et al. developed an association-based method for cleaning up outliers using DBSCAN, k-means, and Apriori techniques to detect outliers, resulting in a smaller MAE and RSME and better results than the decision tree and GBDT algorithms [2].

Most recently, in 2022, F. Antonello et al. introduced a new measure to assess association rules that were retrieved from alarm data for the purpose of identifying FDEPs. This new metric is highly efficient and saves time, unlike conventional association rule metrics. The effectiveness of this new metric was demonstrated through its application to association rules created from databases of simulated alerts and databases of large-scale alarms gathered at CERN's CTI in 2016 [13].

Agrawal, Imielinski, and Swami (AIS). This algorithm generates all the significant association rules between the elements of a large database of customer transactions. Its goal is finding the larger items in the previous| pass for each new transaction [1].

Apriori algorithm. Apriori is used for the frequent exploration of sets of elements and the learning of association rules on relational databases. To do so the algorithm mentioned beforehand, starts by firstly determining the most common individual elements in the database. Secondly, by expanding them to an enormous set of elements. As long as these sets of elements appear sufficiently often in the database [3].

AprioriTID algorithm. AprioriTID considered as a variant of our third algorithm. It is an algorithm that discovers frequent objects in a transaction database. It was proposed in the same article as Apriori, but the difference is that AproriTID uses a different mechanism for counting items support [3].

SET-oriented Mining of Association rules (SETM). The second algorithm takes into account the support for individual items. It also identifies which items are important or frequent in the database, preserving the TID transaction symbol of the generation group share, then it extends the large itemsets of the previous pass in order to generate the candidate itemsets [14].

Direct Hashing and Pruning (DHP). The DHP method developed by M. Houtsma et al. [15] is utilized to boost the Apriori algorithm's performance. To achieve this goal, it uses the hash function to minimize the size of the candidate item set.

FP-Growth. J. Han et al. [16] developed the FP-Growth algorithm, that lets its users find sets of frequent items without using candidate generations. The essence of this method is the use of a special data structure called Frequent Pattern Tree (FP-Tree), which holds the association information of the element sets.

CBAR algorithm. To aid in the finding of big datasets, the CBAR method generates cluster tables. It only requires one database analysis, followed by contrasts with the partial cluster tables; all of this guarantees the accuracy of the results extraction [17].

3. New method for identifying item associations in large-scale distribution

The new approach for identifying item associations in large-scale distribution incorporates data from customers who are engaged in loyalty programs, enabling a more comprehensive understanding of customer behavior and purchasing patterns.

To analyze the relationships between products within the same row or aisle, as well as between different aisles, we use association rule mining techniques. This allows identification of hidden associations between products and better understanding of the factors that influence product preservation.

The hybrid approach combines traditional data analysis with insights gained from customer transactions, resulting in a more accurate representation of real-world customer behavior. Large-scale distributions can use this information to make more informed decisions about product placement and optimize their purchasing strategies.

By focusing on preservation conditions and relationships between products and their placement in aisles, valuable insights are provided into the factors that impact product preservation. This can help large-scale distributions reduce waste and improve the efficiency of their operations.

Overall, the new method for identifying item associations in large-scale distribution offers a unique perspective and added value compared to traditional approaches. Utilizing data from loyalty programs and taking into account the preservation conditions of products aims to contribute to the development of more effective and efficient merchandising practices in the retail industry.

4. Results and discussion

4.1. Dataset

The dataset is a relational collection of files that track consumer orders over time [18]. It includes over 3 million orders from over 200 000 people. We have the sequence of the items purchased at each order for each user, together with the week and time of day the order was placed, as well as a relative estima

Fig. 1. Data model.

We merged order products train and order products prior as product orders for analysis. We integrated this data with orders products based on order id, product id, aisle id, and department id and built the following database with over 30 million rows and 15 columns.

Each row represents an item in a customer's order, and each column provides information about the item, as described below (see Table 1):

Order_id (oid): The unique identifier for the customer's order.
Product_id (pid): The unique identifier for the product.

Add_to_cart_order (cart): The order in which the item was added to the customer's cart.
Reordered (reorder): A binary indicator (0 or 1) that indicates whether the item has been ordered by the customer before.
Product_name (pname): The name of the product.
Aisle_id (aisle_id): The unique identifier for the aisle where the product is located in the store.
Department_id (dept_id): The unique identifier for the department where the product is located in the store.
Aisle (aisle): The name of the aisle where the product is located in the store.
Department (dept): The name of the department where the product is located in the store.
User_id (uid): The unique identifier for the customer.
Eval set (eval): A categorical variable (train, test, or prior) that indicates whether the item is part of the customer's training set, test set, or prior purchases.
Order_number (onum): The order number in which the customer placed the order.
Order_dow (dow): The day of the week on which the customer placed the order (0 = Saturday, 1 = Sunday, etc.).
Order_hof (hof): The hour of the day at which the customer placed the order.
Days_since_prior_order (dspo): The number of days since the customer's previous order.

4.2. Apriori algorithm

The Apriori algorithm is a popular algorithm used in association rule mining to discover interesting relationships between items in a transactional database. The algorithm works by finding frequent itemsets, i.e., sets of items that occur together frequently in the database [3].

Algorithm. The algorithm [3] consists of two main phases:

1. **Support Counting:** In this phase, the algorithm scans the database to count the frequency of each item. The frequent items, i.e., items whose frequency is above a user-defined threshold, are retained for the next phase. The algorithm then generates candidate itemsets of size two, i.e., pairs of frequent items, and counts their frequency in the database. Again, only the frequent itemsets are retained for the next phase. This process is repeated until no new frequent itemsets are found.
2. **Rule Generation:** In this phase, the algorithm generates association rules from the frequent itemsets. An association rule is a statement of the form $A \to B$, where A and B are sets of items. The rule states that if a transaction contains all the items in A, then it is likely to contain all the items in B as well. The algorithm generates rules with a minimum confidence threshold, which is a measure of the reliability of the rule. The confidence of a rule $A \to B$ is defined as:
$$\mathrm{conf}(A \to B) = \frac{\mathrm{support}(A \cup B)}{\mathrm{support}(A)},$$
where $\mathrm{support}(A \cup B)$ is the frequency of the itemset $A \cup B$ in the database, and $\mathrm{support}(A)$ is the frequency of the itemset A in the database. The confidence measures the percentage of transactions containing all the items in A that also contain all the items in B.
The algorithm also computes the lift of each rule, which is a measure of the strength of the association between A and B. The lift of a rule $A \to B$ is defined as:
$$\mathrm{lift}(A \to B) = \frac{\mathrm{support}(A \cup B)}{\mathrm{support}(A) \times support(B)}.$$
The lift measures the ratio of the observed support of A and B to the expected support if A and B were independent. If the lift is greater than 1, it indicates a positive association between A and B.

Input and output. The input to the algorithm is a database of transactions, where each transaction is a set of items. The output of the algorithm is a set of association rules that meet the user-defined minimum support and confidence thresholds. These rules can be used to discover inter-

Table 1. Overview of the Entire Dataset: 30 Million Rows and 15 Columns.

	oid	pid	cart	reorder	pname	aisle_id	dept_id	aisle	dept	uid	eval	onum	dow	hof	dspo
0	2	33120	1	1	Organic Egg Whites	86	16	eggs	dairy eggs	202279	prior	3	5	9	8.0
1	2	28985	2	1	Michigan Organic Kale	83	4	fresh vegeta-bles	produce	202279	prior	3	5	9	8.0
2	2	9327	3	0	Garlic Powder	104	13	spices season-ings	pantry	202279	prior	3	5	9	8.0
3	2	45918	4	1	Coconut Butter	19	13	oils vinegars	pantry	202279	prior	3	5	9	8.0
4	2	30035	5	0	Natural Sweet-ener	17	13	baking ingredi-ents	pantry	202279	prior	3	5	9	8.0
...															
33819104	3421070	16953	2	1	Creamy Peanut Butter	88	13	spreads	pantry	139822	train	15	6	10	8.0
33819105	3421070	4724	3	1	Broccoli Florettes	32	4	packaged produce	produce	139822	train	15	6	10	8.0

esting relationships between items in the database, such as "people who buy diapers are also likely to buy beer" [3].

Deciding Association. To determine if two items A and B are associated, the Apriori algorithm considers the support, confidence, and lift of the rule $A \to B$. A high support value for the itemset $A \cup B$ indicates that the combination of A and B occurs frequently in the database. A high confidence value for the rule $A \to B$ indicates that the probability of B being purchased given that A is purchased is high. In addition to support and confidence, the lift value of the rule is also considered. If the lift value of $A \to B$ is greater than 1, it indicates a positive correlation between the purchase of A and B. If it is less than 1, it indicates a negative correlation, and if it is equal to 1, it indicates no correlation [3].

Results. The aim of the work is to assist supermarkets in improving merchandising, product placement, and the overall customer experience in order to increase satisfaction, loyalty, and ultimately, the return on investment. By taking into account the different conditions of conservation of the products being analyzed, the work offers added value and a unique perspective on the associations between these products. In particular, the focus is on the associations between products that are located in the same rows or aisles, rather than looking at more general associations between all products in the store. By analyzing these more specific relationships, the goal is to identify patterns and trends that can inform strategies for improving merchandising, product placement, and the customer experience in supermarkets. Overall, the aim is to provide valuable insights and recommendations that can help supermarkets to optimize their operations and improve their performance. By focusing on the conditions of conservation of products and the associations between products located in the same rows, this work aims to contribute to the development of more effective and efficient merchandising practices in the retail industry.

Table 2. Association Rules between products of the same aisle.

	Product1	Product2	Support	Confidence	Lift
0	Organic Baby Spinach	Bag of Organic Bananas	0.007	0.133	3.704
1	Organic Hass Avocado	Bag of Organic Bananas	0.009	0.164	5.197
2	Organic Raspberries	Bag of Organic Bananas	0.006	0.107	5.255
3	Organic Strawberries	Bag of Organic Bananas	0.009	0.164	4.161
4	Cucumber Kirby	Banana	0.005	0.067	4.692
5	Large Lemon	Banana	0.006	0.088	3.813
6	Banana	Limes	0.005	0.067	3.216
7	Organic Avocado	Banana	0.008	0.113	4.301
8	Organic Baby Spinach	Banana	0.008	0.109	3.022
9	Organic Fuji Apple	Banana	0.005	0.072	5.392
10	Organic Hass Avocado	Banana	0.005	0.065	2.075
11	Organic Strawberries	Banana	0.008	0.119	3.016
12	Strawberries	Banana	0.006	0.088	4.117
13	Organic Baby Spinach	Organic Avocado	0.005	0.175	4.874
14	Organic Baby Spinach	Organic Hass Avocado	0.005	0.144	4.552
15	Organic Baby Spinach	Organic Strawberries	0.006	0.158	4.014
16	Organic Hass Avocado	Organic Strawberries	0.006	0.192	4.868
17	Organic Raspberries	Organic Strawberries	0.005	0.249	6.329

We used the apriori algorithm of the Mlxtend python library (machine learning extensions) to find related items in the same aisle, which resulted in the following relations.

Moving on to a more general case, consider the relationships between the various aisles of our supermarket's same department.

Finally, we were able to develop relationships between departments.

Table 3. Association Rules between the aisles.

	Aisle1	Aisle2	Support	Confidence	Lift
0	milk	eggs	0.010	0.349	6.734
1	yogurt	eggs	0.011	0.363	6.480
2	fresh herbs	fresh fruits	0.015	0.126	6.320
3	fresh fruits	fresh vegetables	0.068	0.571	6.039
4	fresh fruits	packaged vegetables fruits	0.058	0.487	6.220
5	fresh herbs	fresh vegetables	0.017	0.845	8.929
6	fresh herbs	packaged vegetables fruits	0.011	0.563	7.192
7	packaged vegetables fruits	fresh vegetables	0.050	0.529	6.764
8	milk	packaged cheese	0.017	0.321	6.549
9	yogurt	milk	0.020	0.391	6.981
10	yogurt	packaged cheese	0.019	0.380	6.789
11	yogurt	soy lactosefree	0.012	0.340	6.072
12	fresh herbs	fresh fruits	0.013	0.111	6.592
13	fresh fruits	packaged vegetables fruits	0.040	0.336	6.717
14	fresh herbs	packaged vegetables fruits	0.010	0.503	10.044

Table 4. Association rules between the departments.

	department1	department2	Support	Confidence	Lift
0	babies	beverages	0.011	0.407	2.079
1	babies	dairy eggs	0.014	0.583	2.191
2	babies	deli	0.011	0.606	2.242
3	babies	dry goods pasta	0.011	0.207	2.156
4	babies	produce	0.012	0.574	2.289
...					
105	produce	snacks	0.014	0.094	2.350
106	snacks	bakery	0.012	0.150	2.015
107	snacks	frozen	0.015	0.184	2.566
108	snacks	household	0.018	0.126	2.412
109	snacks	pantry	0.015	0.393	2.673

4.3. Discussion

In recent years, many corporations have been optimizing their resources and minimizing waste by identifying partnerships between items based on various characteristics. While similar items are often associated with each other, our approach in this work is to also consider the association between items and aisles in large-scale distribution settings.

Our analysis aims to identify broad correlations between items to inform strategies for selling or promoting these items together, thereby increasing sales and improving the customer experience by offering unique deals or highlighting complementary products. We also consider the specific preservation conditions that impact the products being analyzed, as some products sold alongside each other may not always be compatible with the same preservation conditions. This led us to identify correlations between items and the aisles in which they are located, providing us with a better understanding of the factors that influence product preservation in these settings.

In addition, we analyze the aisles with low or very low purchase rates, which may contain unrelated goods contributing to low returns on investment. To address this issue, we recommend making these aisles more visible to customers or rearranging the products to improve stock management and reduce losses. Our approach builds on previous works that have developed association rules algorithms, such as Apriori and AprioriTid, and used them to carry out market basket analyses that led to better product placement in supermarkets and increased sales and revenue [3, 12].

5. Conclusion

The new approach presented in this work for identifying item associations in large-scale distribution settings is a unique and valuable contribution to the retail industry. By leveraging data from customers engaged in loyalty programs and using association rule mining techniques, this approach offers a more comprehensive understanding of customer behavior and purchasing patterns. Moreover, the focus on preservation conditions and the relationships between products and their placement in aisles provides valuable insights into factors that impact product preservation and can help reduce waste and improve efficiency.

This hybrid approach, combining traditional data analysis with customer transaction insights, offers a more accurate representation of real-world customer behavior and can lead to more informed decisions about product placement and optimization of purchasing strategies.

Looking to the future, this approach has significant potential for improving merchandising practices in the retail industry. By providing retailers with more accurate insights into customer behavior, the approach can help retailers to increase sales, improve the customer experience, and reduce waste. As more retailers adopt this approach and incorporate customer transaction data, the benefits to the industry will become even more apparent. Furthermore, advances in data analytics and machine learning can further improve the accuracy and efficiency of this approach, leading to even more effective merchandising practices in the future.

[1] Agrawal R., Imieliński T., Swami A. Mining association rules between sets of items in large databases. ACM SIGMOD Record. **22** (2), 207–216 (1993).

[2] Kuang H., Qin R., He M., He X., Duan R., Guo C., Meng X. An Association Rules-Based Method for Outliers Cleaning of Measurement Data in the Distribution Network. Frontiers in Energy Research. **9**, 730058 (2021).

[3] Agrawal R., Srikant R. Fast Algorithms for Mining Association Rules in Large Databases. Proceedings of the 20th International Conference on Very Large Data Bases (VLDB '94). 487–499 (1994).

[4] Liu B., Hsu W., Ma Y. Integrating classification and association rule mining. Proceedings of the Fourth International Conference on Knowledge Discovery and Data Mining (KDD '98). 80–86 (1998).

[5] Holt J. D., Chung S.M. Efficient mining of association rules in text databases. Proceedings of the Eighth International Conference on Information and Knowledge Management (CIKM '99). 234–242 (1999).

[6] Chen G., Wei Q. Fuzzy association rules and the extended mining algorithms. Information Sciences. **147** (1–4), 201–228 (2002).

[7] Wu X., Zhang C., Zhang S. Efficient mining of both positive and negative association rules. ACM Transactions on Information Systems. **22** (3), 381–405 (2004).

[8] Yang X. Y., Liu Z., Fu Y. MapReduce as a programming model for association rules algorithm on Hadoop. The 3rd International Conference on Information Sciences and Interaction Sciences. 99–102 (2010).

[9] Fournier-Viger P., Wu C. W., Tseng V. S. Mining Top-K Association Rules. In: Kosseim L., Inkpen D. (eds.) Advances in Artificial Intelligence. Canadian AI 2012. Lecture Notes in Computer Science, vol. 7310 (2012).

[10] Sahoo J., Das A., Goswami A. An efficient approach for mining association rules from high utility itemsets. Expert Systems with Applications. **42** (13), 5754–5778 (2015).

[11] Wu J. M.-T., Zhan J., Chobe S. Mining Association rules for Low-Frequency itemsets. PLOS ONE. **13** (7), e0198066 (2018).

[12] Unvan Y. Market basket analysis with association rules. Communications in Statistics – Theory and Methods. **50** (7), 1615–1628 (2020).

[13] Antonello F., Baraldi P., Zio E., et al. A Novel Metric to Evaluate the Association Rules for Identification of Functional Dependencies in Complex Technical Infrastructures. Environment Systems and Decisions. **42**, 436–449 (2022).

[14] Houtsma M., Swami A. Set-oriented mining for association rules in relational databases. Proceedings of the Eleventh International Conference on Data Engineering. 25–33 (1995).

[15] Park J. S., Chen M. S., Yu P. S. An effective hash-based algorithm for mining association rules. SIGMOD Rec. **24** (2), 175–186 (1995).

[16] Han J., Pei J., Yin Y. Mining frequent patterns without candidate generation. Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data (SIGMOD '00). 1–12 (2000).

[17] Tsay Y.-J., Chiang J.-Y. CBAR: an efficient method for mining association rules. Knowledge-Based Systems. **18** (2–3), 99–105 (2005).

[18] Instacart Market Basket Analysis.
`https://www.kaggle.com/competitions/instacart-market-basket-analysis/data`.

Революція послуг супермаркетів за допомогою аналізу правил ієрархічної асоціації

Мефтах М.[1], Оунасер С.[1], Ардчір С.[2], Ель Газуані М.[1], Аззуазі М.[1]

[1] *Лабораторія інформаційних технологій та моделювання, Університет Хасана II, Факультет наук Бен М'Сік, Касабланка, Марокко*
[2] *Університет Хасана II, Національна школа комерції та менеджменту, Касабланка, Марокко*

Використання методів аналізу правил асоціації стало центром уваги багатьох дослідників, які прагнуть краще зрозуміти поведінку споживачів. Аналізуючи взаємозв'язки між продуктами та їх розміщенням у проходах, можна отримати цінну інформацію про фактори, які впливають на збереження продуктів у великомасштабних середовищах розповсюдження. Цей підхід може покращити процеси прийняття рішень і оптимізувати результати збереження продуктів, незважаючи на деякі обмеження в якості доступних даних. Крім того, було прийнято гібридний підхід включення транзакцій від клієнтів, які беруть участь у програмі лояльності, щоб заохотити широкомасштабні розповсюдження та краще зрозуміти поведінку клієнтів і покращити їхні стратегії купівлі. Метою цього дослідження є сприяння узгодженості між реальним і віртуальним уявленнями про поведінку клієнтів, що в кінцевому підсумку призведе до покращення результатів купівлі для великомасштабних дистрибуцій.

Ключові слова: *збереження продукції; гібридний підхід; програма лояльності; правила асоціації; апріорі; широкомасштабне поширення.*