|  |  |  |
| --- | --- | --- |
| **id** | **cite** | **abstract** |
| 1 | @article{Kuzior2022GlobalDC,  title={Global Digital Convergence: Impact of Cybersecurity, Business Transparency, Economic Transformation, and AML Efficiency},  author={Aleksandra Kuzior and Tetiana Vasylieva and Olha V. Kuzmenko and Vitaliia Koibichuk and Paulina Brożek},  journal={Journal of Open Innovation: Technology, Market, and Complexity},  year={2022},  url={https://api.semanticscholar.org/CorpusID:253291619}  } | **Global Digital Convergence: Impact of Cybersecurity, Business Transparency, Economic Transformation, and AML Efficiency**  The article substantiates the existence of convergence processes in the field of digitization of countries, taking into account the number of Internet users; people with advanced skills; and indicators of infrastructure (network coverage, population covered by at least a 3G mobile network, population covered by at least a 4G mobile network), access (access to ICT at home, active mobile broadband subscriptions, fixed broadband subscriptions), enablers (fixed broadband over 10 Mbps, mobile data and voice basket, high consumption) and barriers (improved broadband access from 256 kbps to 2 Mbps and from 2 Mbps to 10 Mbps mobile data and voice basket, low consumption) of digital development. The methodological basis for determining the sigma convergence of digitization processes is the coefficient of variation. The values of the coefficient of variation confirmed the high level of convergence between the studied countries in terms of the degree of use of the Internet for conducting digital transactions. The developed econometric model, which describes the influence of statistically significant integral indicators of the national cybersecurity level, ease of doing business, and the anti-money laundering index on the country’s digital development level, made it possible to determine the average trend of dependence on the level of digital development. One hundred four countries were considered for the analysis. The conducted study of the impact of digitalization on economic transformations based on developed quantile regressions made it possible to analyze exactly how the level of digital development for countries with a high level of digitalization and for countries with a low level of digitalization development depends on the value of the national cybersecurity indicator and the ease of doing business, and which countries have the least resistance to the risk factors of their involvement in fraudulent schemes for the purpose of legalizing criminal income. |
| 2 | @article{Gupta2023ECommerceMB,  title={E-Commerce Market Basket Analysis using Apriori Algorithm},  author={Khushi Gupta and Kashyapi Shah and Ameya A Kadam},  journal={INTERANTIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT},  year={2023},  url={https://api.semanticscholar.org/CorpusID:263659317}  } | **E-Commerce Market Basket Analysis using Apriori Algorithm**  This paper presents the usage of the Apriori Algorithm to implement market basket analysis to identify purchase patterns for items that are frequently bought together by customers. The results of this analysis are primarily used to improve sales of multi-product stores by enhancing product placement based on consumers’ shopping habit. In this particular scenario, we have used the data from an online E- commerce store that caters to customers across the world, but primarily focused to the United Kingdom. Key Words: Market Basket, Apriori Algorithm, Association Rules, E-Commerce, Consumer Behaviour, Data Mining |
| 3 | @article{Nafi2023MarketBA,  title={Market Basket Analysis for Sales Transaction in Shopping Stores},  author={Mohd Noor Azam Nafi and Azni Sharlina Zakaria and Nur Izzati Mohamad Arif and Siti Nurhafizah Mohd Shafie and Nasuhar Ab. Aziz and Omar Kairan},  journal={International Journal of Academic Research in Business and Social Sciences},  year={2023},  url={https://api.semanticscholar.org/CorpusID:257119151}  } | **Market Basket Analysis for Sales Transaction in Shopping Stores**  Market Basket Analysis (MBA) system is a widely used technique among marketers, especially for undirected data mining analysis. MBA is also known as product association analysis and the outcome of this analysis is called association rules. The outcome can be used to schedule marketing or advertising strategies and design catalogs for different shop layouts. Discovering the pattern from the customer's buying habits in the shopping stores was collected in their buying transaction. This study aims to compare the item purchased by the respondents between Store A and Store B and to find out the most potential products that customers have bought along with a specific category of products. Convenience non-probability sampling was involved with structured questionnaires of items in store was collected to analyze data. Association analysis was used by analyzing the result from support, confidence, and lift. The findings showed that there are 13 interesting rules of association revealed in this study. Moreover, the result also found that most products that were purchased together are tissues, condiments, instant food, cooking oil, meat, biscuits, dry goods, beverages, and cleaning products. |
| 4 | @article{nvan2020MarketBA,  title={Market basket analysis with association rules},  author={Y{\"u}ksel Akay {\"U}nvan},  journal={Communications in Statistics - Theory and Methods},  year={2020},  volume={50},  pages={1615 - 1628},  url={https://api.semanticscholar.org/CorpusID:213435113}  } | **Market basket analysis with association rules**  This study was conducted in order to make a Market Basket Analysis by using Association Rules. The data used in the study are the sales data of any supermarket received from the Vancouver Island University website. Data were analyzed in the Weka program using a data set containing 225 different products. Apriori and FP Growth, which are Association Rules algorithms, were tried in order. Since the data set is categorical, the Apriori algorithm did not yield any results. Therefore, the FP Growth algorithm was used and the top 10 rules were given according to the conviction value. The best rule accordingly; a customer who buys Milk, Sweet Relish and Pepperoni Pizza (Frozen) also gets eggs. Best rule with 21.06 Conviction and 1 (100%) confidence values are this rule. 24 customers who received these 3 products in the dataset received eggs. Similarly, also other rules were interpreted in this study. As a result, product placement in the supermarket can be made according to these rules. Thus, sales of these products will increase and supermarket revenue will increase directly. |
| 5 | @article{Priyanto2022IMPLEMENTATIONOM,  title={IMPLEMENTATION OF MARKET BASKET ANALYSIS WITH APRIORI ALGORITHM IN MINIMARKET},  author={Abdul Hafiidh Priyanto and Amalia Beladinna Arifa},  journal={Jurnal Teknik Informatika (Jutif)},  year={2022},  url={https://api.semanticscholar.org/CorpusID:256128251}  } | **IMPLEMENTATION OF MARKET BASKET ANALYSIS WITH APRIORI ALGORITHM IN MINIMARKET**  The rapid growth of the retail business has an impact on increasing the economic growth of the community. The retail business has high profit potential in areas that have a large population such as Indonesia. A retail business that is popular among the public is a modern market retail business or convenience store. With the rapid growth, it gives a tendency between convenience stores to compete. By designing a marketing strategy is one of the efforts to win the competition in supermarkets. Management needs to understand the purchase behavior made by customers, this action is useful to find out the products that customers are popularly buying. Association algorithm is a form of algorithm in the field of data mining that serves to provide correlation between one item and another. there are several popular algorithms in applying association algorithms one of which is the a priori algorithm created by Agrawal and Srikant in 1994. To support the understanding of customer purchase patterns, it is necessary to implement market basket analysis that has the ability to recognize pattern patterns from transaction data in a convenience store. Performance in market basket analysis also needs to be tested to handle a lot of transaction data, considering that the recording of sales transaction data continues to run over time. The implementation carried out using flask is one of the implementations that is relevant to technological developments, this implementation results in a relatively short data speed with the factor that the magnitude of transaction data is middle to lower, which is 14,963 transaction data. |
| 6 | @article{Nuraeni2022OPTIMIZATIONOM,  title={OPTIMIZATION OF MARKET BASKET ANALYSIS USING CENTROID-BASED CLUSTERING ALGORITHM AND FP-GROWTH ALGORITHM},  author={Fitri Nuraeni and Dewi Tresnawati and Yoga Handoko Agustin and Gisna Fauzi},  journal={Jurnal Teknik Informatika (Jutif)},  year={2022},  url={https://api.semanticscholar.org/CorpusID:256127322}  } | **OPTIMIZATION OF MARKET BASKET ANALYSIS USING CENTROID-BASED CLUSTERING ALGORITHM AND FP-GROWTH ALGORITHM**  The proliferation of the food and beverage sales business requires the creativity of business owners to offer their flagship products to every consumer, both new and subscribed consumers. A large number of menu choices makes the ordering process long because consumers are confused about which menu will be the best choice. the seller to be able to provide the right recommendations so that orders can take place faster. Shopping cart analysis is an activity that has often been done to find out the items found that are sold simultaneously. The FP-Growth association method is a faster algorithm for generating association rules, but the association process in large dataset sizes tends to add large items so that the accuracy value of association rules decreases. So that in this study, the grouping of datasets was carried out using a clustering model with a centroid-based algorithm, namely k-means, k-medoids, and fuzzy c-means. This research was conducted through dataset collection, dataset preparation, clustering modeling, evaluation of clustering models using DBI and silhouette index, association modeling, and evaluation of association models using lift ratio. The results of this study showed that the clustering model with the best DBI and silhouette index values ​​was at k=3 for k-means, k=2 for k-medoids, and k=7 for fuzzy c-means. The number of association rules is generated from the grouped data set using fuzzy c-means, but the highest average lift ratio is in the association rules generated from the grouping data set using k-means. From the association model using k-means and FP-Growth, 32 unique association rules were found with the 4 most frequently found items, namely cireng chili oil, regal milk coffee, banana cheese, and vietnam drip. |
| 7 | @article{Karnila2022MARKETBA,  title={MARKET BASKET ANALYSIS ON TRANSACTION DATA USING THE APRIORI ALGORITHM},  author={Sri Karnila and Akbar Rizkyandi and Rio Kurniawan and Nurjoko Nurjoko},  journal={Jurnal TAM (Technology Acceptance Model)},  year={2022},  url={https://api.semanticscholar.org/CorpusID:252757734}  } | **MARKET BASKET ANALYSIS ON TRANSACTION DATA USING THE APRIORI ALGORITHM**  This research aims to get information about the relationship between sales patterns carried out by CV. Dian Abadi Jaya workshop by using APRIORI algorithms through transaction data sets carried out by customers. The subject of research is a record of shopping cart transactions made by customers, namely vehicle parts sales transactions and vehicle repair service transactions. The data collection techniques used are interviews and documentation. The criteria used in this research are a minimum of frequent itemset of 20 transactions with support criteria of 1,7%, confidence value of 40% and lift ratio value above 1. The results of the research have produced 9 sales pattern relationships with the highest confidence of 100%. The results that have been obtained are expected to help the CV. Dian Abadi Jaya workshop in making a decision for the next sale. |
| 8 | @article{Patwary2021MarketBA,  title={Market Basket Analysis Approach to Machine Learning},  author={Abu Hasnat Patwary and Md Tamim Eshan and Prazzal Debnath and Abdus Sattar},  journal={2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)},  year={2021},  pages={1-9},  url={https://api.semanticscholar.org/CorpusID:241592133}  } | **Market Basket Analysis Approach to Machine Learning**  Market Basket Analysis is an important aspect of a retail organization's analytical framework for deciding where products should be placed and developing sales promotions for various segments of consumers to increase customer loyalty and, as a result, benefit. The market is the well-known activity of ARM ultimately used for business intelligent decisions. Market Basket Analysis is a data mining technique that can be used in various fields, such as marketing, bioinformatics, education field, nuclear science, etc. Our objective here is that traders can further improve their business. At the same time, they can make more profits and invest more and more. This will be of great benefit. If the business of traders increases, many people will get employment, as well as customers, will be able to make their way of life easier by getting familiar with different things. Here is a discussion on their methods so that traders in Bangladesh can make proper investments in the right place. The frequent itemsets are mined from the database using the Apriori algorithm and then the association rules are generated. The project will assist supermarket managers in determining the relationship between the items that their customers purchase. |
| 9 | @article{Sussolaikah2021MarketBA,  title={Market Basket Analysis for Determination of Consumer Behavior at XYZ Stores Using R Programming},  author={Kelik Sussolaikah},  journal={Advance Sustainable Science, Engineering and Technology},  year={2021},  url={https://api.semanticscholar.org/CorpusID:243842387}  } | **Market Basket Analysis for Determination of Consumer Behavior at XYZ Stores Using R Programming** Data mining is one of the fields of science in the world of informatics which has an important role, especially with regard to data. There are many algorithms and methods that can be used to process data. The paper this time the author tries to conduct research on consumer behavior by using one of the data mining techniques, namely market basket analysis. This research uses the R Programming tool, where it is hoped that the research can be carried out effectively and efficiently. Based on the research conducted, it is known that there has been a significant purchase of several items that have been described as a plot. The tendency of consumers to buy several items followed by other items can be a consideration for arranging the layout of goods on the sales shelf or arranging product stock in a supermarket. |
| 10 | @article{MarcelinoIrawan2021MARKETBA,  title={MARKET BASKET ANALYSIS METHOD ON SALES DATA USING FP-GROWTH ALGORITHM},  author={Kenny Marcelino Irawan and Tina Tri Wulansari and Nariza Wanti Wulan Sari},  journal={MULTICA SCIENCE AND TECHNOLOGY (MST) JOURNAL},  year={2021},  url={https://api.semanticscholar.org/CorpusID:259808321}  } | **MARKET BASKET ANALYSIS METHOD ON SALES DATA USING FP-GROWTH ALGORITHM**  Product promotion is a way for business owners to increase sales of existing goods. Business owners could use association rules as a consideration of product promotion policies. Determination of association rules can be determined using the market basket analysis method with the fp-growth algorithm. From the research, there are 248 association rules for goods that are often purchased simultaneously with minimum support of 0.01 and minimum confidence of 0.1. Of the 248 rules, there are seven rules that have a confidence value of more than 0.5. Of the seven rules, flour appeared five times, and the cake mat and cake box had the highest confidence value of 0.62 and a lift of 5.54. Therefore, we recommend shop owners to place wheat flour on the main display close to items often purchased together, such as sweetened condensed milk, sugar, powdered margarine, and margarine. In addition, shop owners can also promote bundling cake boxes and cake mats. |
| 11 | @article{Valle2021FindingHS,  title={Finding Hierarchical Structures of Disordered Systems: An Application for Market Basket Analysis},  author={Mauricio A. Valle and Gonzalo A. Ruz},  journal={IEEE Access},  year={2021},  volume={9},  pages={1626-1641},  url={https://api.semanticscholar.org/CorpusID:230999145}  } | **Finding Hierarchical Structures of Disordered Systems: An Application for Market Basket Analysis**  Complex systems can be characterized by their level of order or disorder. An ordered system is related to the presence of system properties that are correlated with each other. For example, it has been found in crisis periods that the financial systems tend to be synchronized, and symmetry appears in financial assets’ behavior. In retail, the collective purchasing behavior tends to be highly disorderly, with a diversity of correlation patterns appearing between the available market supply. In those cases, it is essential to understand the hierarchical structures underlying these systems. For the latter, community detection techniques have been developed to find similar behavior clusters according to some similarity measure. However, these techniques do not consider the inherent interactions between the multitude of system elements. This paper proposes and tests an approach that incorporates a hierarchical grouping process capable of dealing with complete weighted networks. Experiments show that the proposal is superior in terms of the ability to find minimal energy clusters. These minimum energy clusters are equivalent to system states (market baskets) with a higher probability of occurrence; therefore, they are interesting for marketing and promotion activities in retail environments. |
| 12 | @inproceedings{Sinha2021ImplyingAR,  title={Implying Association Rule Mining and Market Basket Analysis for Knowing Consumer Behavior and Buying Pattern in Lockdown - A Data Mining Approach},  author={Anurag Sinha},  year={2021},  url={https://api.semanticscholar.org/CorpusID:235559451}  } | **Implying Association Rule Mining and Market Basket Analysis for Knowing Consumer Behavior and Buying Pattern in Lockdown - A Data Mining Approach**  Buyer practices have changed as individuals are figuring out how to live with the new truth of COVID-19. Take-out and conveyance orders have expanded, and our customer has added new items to their menu because of new client inclinations. With every one of the continuous changes, the customer had numerous unanswered inquiries, for example, Smartbridge has broad involvement with café innovation development Café TECHNOLOGY CAPABILITIES :Are the most famous items as yet unchanged after COVID? :Which are the most sold item blends now? :What is the acknowledgment of new things? :What are clients purchasing alongside new things? :How have liquor deals changed? The customer previously had reports that followed item deals and operational measurements, notwithstanding, there was a need to get a more profound knowledge into item examination. The customer expected to recognize what items and introductions were being sold all the more frequently, measure the acknowledgment of new items, and figure out what items clients buy together to improve advertising efforts, advancements, and deals. he E-business industry is filling immensely in the Indian market. The modest 4G web bundles in India clearly gives a push to these ventures. Thus, as Covid19 first hit in Quite a while, individuals got terrified to go out from their homes in light of the fact that, in their mind, it's a dread of Covid. They even wonder whether or not to go out to purchase fundamental (FMCG) products. Frenzy purchasing additionally has seen and to stay away from this dread of COVID-19, individuals are offering inclinations to the E-Commerce destinations to purchase fundamental products and a few clients are new which joined to purchase fundamental merchandise during this Pandemic Lockdown period. Numerous clients are moving their purchasing conduct from disconnected retail locations to online stores. This paper examines the customer buying pattern during lockdown. |
| 13 | @inproceedings{Pfutzenreuter2021MACHINELF,  title={MACHINE LEARNING FOR SALES INSIGHTS WITH ASSOCIATION RULES MARKET BASKET ANALYSIS},  author={Thais Carreira Pfutzenreuter and Edson Pinheiro de Lima},  year={2021},  url={https://api.semanticscholar.org/CorpusID:240416425}  } | **MACHINE LEARNING FOR SALES INSIGHTS WITH ASSOCIATION RULES MARKET BASKET ANALYSIS**  With an extensive variety of data collection, companies have been investing in mining association rules from their databases. The detection of association relationships between large quantities of business transactions can help in marketing, financial achievements and other business strategic fields. The purpose of this study is to investigate association rules between an American food company products’ sales transaction. In order to follow machine learning best practices, a descriptive analysis was applied before Apriori algorithm to get some initial insights of customer behavior. Descriptive analysis revealed bananas and bag of organic bananas as the most frequent products and that a considerable number of customers buy their orders weekly or monthly. Apriori algorithm results identified the best rule as: once buying organic raspberries, the customer has three more chances to also buy organic strawberries. Additionally, the rule with the highest confidence was organic fuji apple with bananas, which means that 38% of all the transactions clients that buy fuji apples also buy bananas. The present study strengthens the hypothesis that association rules can provide strategic knowledge, since Apriori algorithm diagnosed meaningful tendencies for marketing decisions. The dataset was provided by the company for Kaggle competitors. However, it is important to highlight our focus consists on introducing an academic investigation for preliminary insights and not on winning competitions. |
| 14 | @article{Yustiana2021BASKETMA,  title={BASKET MARKET ANALYSIS USING R-BASED APRIORI ALGORITHM TO FIND INFORMATION FROM SALES DATA},  author={Indra Yustiana and Irvan Sutarha and Nindia Maulawati and Ilham Maulana Yusuf},  journal={INTERNATIONAL JOURNAL ENGINEERING AND APPLIED TECHNOLOGY (IJEAT)},  year={2021},  url={https://api.semanticscholar.org/CorpusID:251782431}  } | **BASKET MARKET ANALYSIS USING R-BASED APRIORI ALGORITHM TO FIND INFORMATION FROM SALES DATA**  Market Basket Analysis is a data mining technique that is used to determine which products a customer will buy simultaneously by analyzing a list of customer transactions. By knowing these products, an e-commerce system can create or develop a customer profile system and can determine its own customer catalog layout. This journal discusses data mining techniques, with association rules that can help check customer buying behavior and increase sales. The result can provide reference prices for cross selling, designing promotions and placing merchandise in stores increasing sales |
| 15 | @article{HS2023ASO,  title={A STUDY ON SOCIAL MEDIA STRATEGIES FOR ONLINE GROCERY SHOPPING CART AT BIG BASKET BANGALORE},  author={Venkatehagowda H S and Dr. Naveen G},  journal={INTERANTIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT},  year={2023},  url={https://api.semanticscholar.org/CorpusID:265412817}  } | **A STUDY ON SOCIAL MEDIA STRATEGIES FOR ONLINE GROCERY SHOPPING CART AT BIG BASKET BANGALORE**  Online grocery spending has been increasingly popular in current years as a outcome of the expansion of e- commerce, which has completely changed how customers interact with retail platforms. Businesses are looking for imaginative ways to improve their social media occurrence and interconnect with customers as the online grocery industry gets more competitive. This research explores the world of social media strategies designed especially for Big Basket's online grocery store, a significant player in the online grocery market. |
| 16 | @article{Dwiputra2023EvaluatingTP,  title={Evaluating the Performance of Association Rules in Apriori and FP-Growth Algorithms: Market Basket Analysis to Discover Rules of Item Combinations},  author={Dedy Dharmadi Cakra Dwiputra and Agung Mulyo Widodo and Habibullah Akbar and Gerry Firmansyah},  journal={Journal of World Science},  year={2023},  url={https://api.semanticscholar.org/CorpusID:261602737}  } | [**Evaluating the Performance of Association Rules in Apriori and FP-Growth Algorithms: Market Basket Analysis to Discover Rules of Item Combinations.**](https://www.semanticscholar.org/paper/Evaluating-the-Performance-of-Association-Rules-in-Dwiputra-Widodo/d08ef132543aa30d5b69b7480dea1f377ee3726e)  This study focuses on applying data mining techniques, especially association rules mining using the Apriori and FP-GROWTH algorithms, for market basket analysis on PT. XYZ is a pharmaceutical company in Indonesia. A quantitative methodology uses a dataset of 100,498 transactions originating from 432,356 rows of data covering July to December 2022 in the JABODETABEK area. Apriori and FP-GROWTH algorithms are applied for association rules mining. The results show that FP-GROWTH has the fastest execution time of 84,655 seconds. However, the memory usage for the Apriori algorithm is the lowest at 482.32 MiB, with increments of: 0.21 MiB. For the rules generated, the two algorithms, both Apriori and FP-GROWTH, produce the same number of rules and values of support, confidence, lift, Bi-Support, Bi-Confidence, and Bi-Lift. In conclusion, Apriori is recommended for sales datasets if memory usage and ease of implementation are important. However, if the speed of execution time and a large amount of data are considered, FP-GROWTH is a better choice because the execution time is faster for large amounts of data. However, the choice of algorithm depends on the specific analysis objectives, itemset size, data scale, and computational capabilities. Results from association rules mining provide evidence of product popularity, purchasing patterns, and opportunities for strategic marketing and inventory management. These findings can help PT. XYZ improves business efficiency, understands customer behavior, and increases profitability. |
| 17 | @article{Wahyuningsih2023COMPARISONOM,  title={COMPARISON OF MARKET BASKET ANALYSIS METHOD USING APRIORI ALGORITHM, FREQUENT PATTERN GROWTH (FP- GROWTH) AND EQUIVALENCE CLASS TRANSFORMATION (ECLAT) (CASE STUDY: SUPERMARKET “X” TRANSACTION DATA FOR 2021)},  author={Rina Wahyuningsih and Agus Suharsono and Nur Iriawan},  journal={Business and Finance Journal},  year={2023},  url={https://api.semanticscholar.org/CorpusID:268149601}  } | **COMPARISON OF MARKET BASKET ANALYSIS METHOD USING APRIORI ALGORITHM, FREQUENT PATTERN GROWTH (FP- GROWTH) AND EQUIVALENCE CLASS TRANSFORMATION (ECLAT) (CASE STUDY: SUPERMARKET “X” TRANSACTION DATA FOR 2021**  The retail industry continues to grow and develop in Indonesia. The retail sector as a provider of goods used in everyday life has long started digital transformation in its business. Digital technology helps the retail industry collect valuable customer data. Business analytic is the use of data, information technology and statistical analysis, which is used to obtain information about a business and make decisions based on facts. Business analytic turns data into steps or actions in the context of making business decisions. Consumer needs and purchasing behavior can be predicted with big data-based technology. Association Rule is a technique in data mining to find the relationship between items in an item set combination. One of the utilizations of the association rule method is market basket analysis. Algorithms that can be used to analyze consumer purchasing patterns include the Apriori algorithm, Frequent Pattern Growth (FP-Growth) which represents a database structure in a horizontal format, and the Equivalence Class Transformation (ECLAT) algorithm which represents a vertical data format. In addition, this research will first analyze the complexity of the algorithm based on the time complexity in running the algorithm. This analysis uses these three algorithms, which are applied to Supermarket "X" transaction data in 2021, namely 136,202 transactions. The measure of goodness that is used to find out the best algorithm uses support and confidence values. The results show that the ECLAT algorithm is the most superior algorithm compared to the others based on the execution time required by the algorithm. The support value used in forming associations in the ECLAT algorithm is 1%, resulting in 19 rules. From the results of these rules, the highest support value was generated by the purchase of Indomie goreng special and Indomie ayam bawang, where as many as 1,362 shopping transactions bought these two items together or 2.71% of the total transactions. |
| 18 | @article{Gupta2023ECommerceMB,  title={E-Commerce Market Basket Analysis using Apriori Algorithm},  author={Khushi Gupta and Kashyapi Shah and Ameya A Kadam},  journal={INTERANTIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT},  year={2023},  url={https://api.semanticscholar.org/CorpusID:263659317}  } | **E-Commerce Market Basket Analysis using Apriori Algorithm**  This paper presents the usage of the Apriori Algorithm to implement market basket analysis to identify purchase patterns for items that are frequently bought together by customers. The results of this analysis are primarily used to improve sales of multi-product stores by enhancing product placement based on consumers’ shopping habit. In this particular scenario, we have used the data from an online E- commerce store that caters to customers across the world, but primarily focused to the United Kingdom. Key Words: Market Basket, Apriori Algorithm, Association Rules, E-Commerce, Consumer Behaviour, Data Mining |
| 19 | @article{Idris2022ComparisonOA,  title={Comparison of Apriori, Apriori-TID and FP-Growth Algorithms in Market Basket Analysis at Grocery Stores},  author={Andi Ilhamsyah Idris and Eliyah Acantha M Sampetoding and Valian Yoga Pudya Ardhana and Irene Maritsa and Adrisumatri Sakri and Hidayatullah Ruslan and Esther Sanda Manapa},  journal={The IJICS (International Journal of Informatics and Computer Science)},  year={2022},  url={https://api.semanticscholar.org/CorpusID:257018537}  } | **Comparison of Apriori, Apriori-TID and FP-Growth Algorithms in Market Basket Analysis at Grocery Stores.**  Market Basket Analysis is an analysis of consumer behavior specifically from a certain group/group. Market Basket Analysis is generally used as a starting point for seeking knowledge from a data transaction when we do not know what specific pattern we are looking for. Market Basket Analysis in this study is applied to the search for patterns of purchasing groceries at grocery stores and then analyzed by season. This study aims to compare the Apriori, Apriori TID and FP-Growth methods in determining consumer transaction behavior and calculating the quantity of consumer transactions in several seasons based on data obtained from the Market Basket Analysis database. In the results of this study, it is known that FP-Growth has the best performance among the other two algorithms, but uses more memory than other algorithms. The Apriori-TID algorithm uses lighter and faster memory than the Apriori Algorithm |
| 20 | @article{Almaslamani2020UsingBD,  title={Using Big Data Analytics to Design an Intelligent Market Basket-Case Study at Sameh Mall},  author={Farah Almaslamani and Raneem Abuhussein and Hanan Saleet and Laith Abuhilal and Nader S. Santarisi},  journal={International journal of engineering research and technology},  year={2020},  volume={13},  pages={3444-3455},  url={https://api.semanticscholar.org/CorpusID:234521068}  } | **[Using Big Data Analytics to Design an Intelligent Market Basket-Case Study at Sameh Mall](https://www.semanticscholar.org/paper/Using-Big-Data-Analytics-to-Design-an-Intelligent-Almaslamani-Abuhussein/c943c8b7029d083bcd02fc58a2d2849aa2e6e4c3)**  The long term social, economic and health impacts of the COVID-19 pandemic are still unknown Retailers should think about the impact this pandemic will have on the customer relationship Another factor that is rigorously influencing the retail industry is the digital transformation With the digital transformation worldwide, coupled with the exponential growth of the use of big data analytics, retailers can use intelligent market basket analysis to help in shoring up customer relationships This study uses big data analytics to design and analyze intelligent market basket in one top retailer in Jordan, "Sameh Mall" It aims to help managers to improve customer relationship while increasing sales Customers' behavioral similarities analysis results in different baskets, which contain items commonly bought together Such baskets are displayed physically in stores and are displayed online as promotions This study results are interesting, enabling Sameh Mall to send recommendations to VIP customers through their account on the online application;and recommendations for physical cross promotional or cross merchandising leading to increases in basket size, increase in sales, as well as increase in customer satisfaction © International Research Publication House |
| 21 | @article{Karnila2022MARKETBA,  title={MARKET BASKET ANALYSIS ON TRANSACTION DATA USING THE APRIORI ALGORITHM},  author={Sri Karnila and Akbar Rizkyandi and Rio Kurniawan and Nurjoko Nurjoko},  journal={Jurnal TAM (Technology Acceptance Model)},  year={2022},  url={https://api.semanticscholar.org/CorpusID:252757734}  } | **[MARKET BASKET ANALYSIS ON TRANSACTION DATA USING THE APRIORI ALGORITHM](https://www.semanticscholar.org/paper/MARKET-BASKET-ANALYSIS-ON-TRANSACTION-DATA-USING-Karnila-Rizkyandi/7226050d5afbc8156bb27e28f19a696f3842ce37)**  This research aims to get information about the relationship between sales patterns carried out by CV. Dian Abadi Jaya workshop by using APRIORI algorithms through transaction data sets carried out by customers. The subject of research is a record of shopping cart transactions made by customers, namely vehicle parts sales transactions and vehicle repair service transactions. The data collection techniques used are interviews and documentation. The criteria used in this research are a minimum of frequent itemset of 20 transactions with support criteria of 1,7%, confidence value of 40% and lift ratio value above 1. The results of the research have produced 9 sales pattern relationships with the highest confidence of 100%. The results that have been obtained are expected to help the CV. Dian Abadi Jaya workshop in making a decision for the next sale. |
| 22 | @inproceedings{Nurmayanti2021MarketBA,  title={Market Basket Analysis with Apriori Algorithm and Frequent Pattern Growth (Fp-Growth) on Outdoor Product Sales Data},  author={Wiwit Pura Nurmayanti and Hanipar Mahyulis Sastriana and Abdul Rahim and Muhammad Munawir Gazali and Ristu Haiban Hirzi and Zuhut Ramdani and Muhammad Malthuf},  year={2021},  url={https://api.semanticscholar.org/CorpusID:233418197}  } | **[Market Basket Analysis with Apriori Algorithm and Frequent Pattern Growth (Fp-Growth) on Outdoor Product Sales Data](https://www.semanticscholar.org/paper/Market-Basket-Analysis-with-Apriori-Algorithm-and-Nurmayanti-Sastriana/5bb633b6af91b980a8383da6bd16d098232022b8)**  Indonesia is an equatorial country that has abundant natural wealth from the seabed to the top of the mountains, the beauty of the country of Indonesia also lies in the mountains that it has in various provinces, for example in the province of West Nusa Tenggara known for its beautiful mountain, namely Rinjani. The increase in outdoor activities has attracted many people to open outdoor shops in the West Nusa Tenggara region. Sales transaction data in outdoor stores can be processed into information that can be profitable for the store itself. Using a market basket analysis method to see the association (rules) between a number of sales attributes. The purpose of this study is to determine the pattern of relationships in the transactions that occur. The data used is the transaction data of outdoor goods. The analysis used is the Association Rules with the Apriori algorithm and the frequent pattern growth (FP-growth) algorithm. The results of this study are formed 10 rules in the Apriori algorithm and 4 rules in the FP-Growth algorithm. The relationship pattern or association rule that is formed is in the item "if a consumer buys a portable stove, it is possible that portable gas will also be purchased" at the strength level of the rules with a minimum support of 0.296 and confidence 0.774 at Apriori and 0.296 and 0.750 at FP-Growth |
| 23 | @inproceedings{Bala2016PerformanceAO,  title={Performance Analysis of Apriori and FP-Growth Algorithms ( Association Rule Mining ) 1},  author={Alhassan Bala and Mansur Zakariyya Shuaibu and Zaharaddeen Karami Lawal},  year={2016},  url={https://api.semanticscholar.org/CorpusID:31009089}  } | **[Performance Analysis of Apriori and FP-Growth Algorithms ( Association Rule Mining ) 1](https://www.semanticscholar.org/paper/Performance-Analysis-of-Apriori-and-FP-Growth-(-)-1-Bala-Shuaibu/4a8c1fd38fd1541d1626c7a20a346d38e66584a4)**  Association rule mining has become popular among marketers and organizations. In fact, an example of association rule mining is referred to as market basket analysis. The task is to find which items are frequently purchased together. This knowledge can be used by professionals to plan where to place items that are frequently bought together closely to each other, thus helping to improve the sales. It involves the relationships between items in a data set. Association rule mining finds out item sets which has minimum support and are represented in a relatively high number of transactions. These transactions are simply known as frequent item sets. The algorithms that use association rules are divided into two stages, first is to find the frequent sets and the second is to use these frequent sets to generate the association rules. In this paper we used Weka to compare two algorithms (Apriori and FP-growth) based on execution time and database scan parameters used are; number of instances, confidence and support levels it is categorically clear that FP-Growth algorithm is better than apriori algorithm. |
| 24 | @article{Gino2023ExploratoryAO,  title={Exploratory Analysis on Market Basket Data using Network Visualization},  author={Henrique L. S. Gino and Diogenes S. Pedro and Jean R. Ponciano and Claudio D. G. Linhares and Agma J. M. Traina},  journal={Anais do XII Brazilian Workshop on Social Network Analysis and Mining (BraSNAM 2023)},  year={2023},  url={https://api.semanticscholar.org/CorpusID:259312945}  } | **[Exploratory Analysis on Market Basket Data using Network Visualization](https://www.semanticscholar.org/paper/Exploratory-Analysis-on-Market-Basket-Data-using-Gino-Pedro/19125a0083ed03fc80d1e5f9651453435600bb42)**  Market basket analysis is a powerful technique for understanding customer behavior and optimizing business strategies based on that understanding. Market-based analysis over time using visualization techniques can provide insights into market trends and relations, simplify complex data, and communicate insights effectively, which can help organizations make more informed decisions. This paper leverages a dataset focused on the users’ incomes and temporal aspects of market purchases. We modeled this dataset as three distinct temporal networks and performed an exploratory evaluation identifying patterns and anomalies in the data. More specifically, we identified groups of related products, indicating thematic purchases, and evaluated the impact of demographic factors, such as income, on customer spending. |
| 24 | @article{Ghassani2021MARKETBA,  title={MARKET BASKET ANALYSIS USING THE FP-GROWTH ALGORITHM TO DETERMINE CROSS-SELLING},  author={Fildzah Zia Ghassani and Asep Jamaludin and Agung Susilo Yuda Irawan},  journal={Jurnal Informatika Polinema},  year={2021},  url={https://api.semanticscholar.org/CorpusID:239700060}  } | **[MARKET BASKET ANALYSIS USING THE FP-GROWTH ALGORITHM TO DETERMINE CROSS-SELLING](https://www.semanticscholar.org/paper/MARKET-BASKET-ANALYSIS-USING-THE-FP-GROWTH-TO-Ghassani-Jamaludin/41a43b2d7dbfd5eb58d547d046cc27556453d194)**  . KAOCHEM Sinergi Mandiri Cooperative is a cooperative that provides various kinds of basic needs such as basic foodstuffs that can meet the needs of its members. The cooperative transaction data is only stored as a report. Association rules are a method in data mining that functions to identify items that have a value that is likely to appear simultaneously with other items. One implementation of the association method is Market Basket Analysis. The data used are transaction data for November 2019. Data mining is one of the processes or stages of the KDD method. The data mining process is carried out using the FP-Growth algorithm, which is one of the algorithms for calculating the sets that often appear from data. Researchers analyzed transaction data using the Rapid Miner Studio tools. In the data mining process using FP-Growth the researcher determines a minimum support value of 3% and a minimum confidence of 50%. The association process using these values ​​produces 3 strong rules, namely if ades 350 ml, then fried / lontong with a support value of 0.030 and confidence 0.556 and if fried st, then fried / lontong with a support value of 0.048 and confidence 0.639, and if nasi uduk / bacang , then fried / rice cake with a support value of 0.031 and confidence 0.824. The results of the association rules can be applied using one of the marketing techniques, namely cross-selling to increase the sales of the cooperative. |
| 25 | @article{Christian2021RealMB,  title={Real Market Basket Analysis using Apriori and Frequent Pattern Tree Algorithm},  author={Michael Albert Christian and Nathanael Nathanael and Annisa Mauliani and Ariani Indrawati and L. Manik and Zaenal Akbar},  journal={Proceedings of the 2021 International Conference on Computer, Control, Informatics and Its Applications},  year={2021},  url={https://api.semanticscholar.org/CorpusID:246801682}  } | **[Real Market Basket Analysis using Apriori and Frequent Pattern Tree Algorithm](https://www.semanticscholar.org/paper/Real-Market-Basket-Analysis-using-Apriori-and-Tree-Christian-Nathanael/b6e08979dec9ac3e8d230b8f5d951dc3edf8a91a)**  Recently, data mining has been implemented in various fields, including business and telecommunications. Data mining is a technique for extracting and detecting patterns in massive data sets that combines machine learning, statistics, and database systems. One of the most important use-cases in data mining is finding the high-frequency patterns between the set of itemset called association rules. Association rule mining is a well-researched technique for finding some relations between variables in large databases. This paper aims to measure the performance of the Apriori and Frequent Pattern Tree algorithms by comparing them using several points of comparison. Then we compared the outputs, whether they produce the same or different rules, to find out whether the way the two algorithms work is similar or not. After that, we looked for the itemsets that best match the reality in the market by giving them to a user who had transaction data from his spare parts shop. |
| 27 | @article{Winarti2023DataMM,  title={Data Mining Modeling Feasibility Patterns of Graduates Ability With Stakeholder Needs Using Apriori Algorithm},  author={Titin Winarti and Henny Indriyawati},  journal={International Journal of Information Technology and Business},  year={2023},  url={https://api.semanticscholar.org/CorpusID:260906545}  } | **[Data Mining Modeling Feasibility Patterns of Graduates Ability With Stakeholder Needs Using Apriori Algorithm](https://www.semanticscholar.org/paper/Data-Mining-Modeling-Feasibility-Patterns-of-With-Winarti-Indriyawati/1f27db863dc94c637731e03b007e2a471dd7bf05)**  This The speed of information, the accuracy of data, the ease of information services, and accountability are very important reasons for the implementation of the system. Semarang University (USM) is a private university in Semarang that has the most 2 students in Central Java. Based on the 2019 USM tracer data showing horizontal alignment, namely how close the relationship between the field of study and alumni work is, it appears that there is still a discrepancy in the ability of graduates with stakeholders.  The Apriori algorithm is the best-known algorithm for finding high-frequency patterns  Rules that state associations between attributes are often called affinity analysis or market basket analysis. The use of the Apriori Algorithm in data mining calculations using data from the Semarang University tracer that the limit of the minimum support is 50% and the minimum confidence is 100% so that it forms 4 rules. From the four rules produced that modeling using the Apriori Algorithm can produce several rule formations so that it can provide an evaluation to the University for compiling steps, this can be seen because the resulting rules are different because each graduate relationship with the desired desires and different styles. |
| 28 | @article{MarcelinoIrawan2021MARKETBA,  title={MARKET BASKET ANALYSIS METHOD ON SALES DATA USING FP-GROWTH ALGORITHM},  author={Kenny Marcelino Irawan and Tina Tri Wulansari and Nariza Wanti Wulan Sari},  journal={MULTICA SCIENCE AND TECHNOLOGY (MST) JOURNAL},  year={2021},  url={https://api.semanticscholar.org/CorpusID:259808321}  } | **MARKET BASKET ANALYSIS METHOD ON SALES DATA USING FP-GROWTH ALGORITHM**  Product promotion is a way for business owners to increase sales of existing goods. Business owners could use association rules as a consideration of product promotion policies. Determination of association rules can be determined using the market basket analysis method with the fp-growth algorithm. From the research, there are 248 association rules for goods that are often purchased simultaneously with minimum support of 0.01 and minimum confidence of 0.1. Of the 248 rules, there are seven rules that have a confidence value of more than 0.5. Of the seven rules, flour appeared five times, and the cake mat and cake box had the highest confidence value of 0.62 and a lift of 5.54. Therefore, we recommend shop owners to place wheat flour on the main display close to items often purchased together, such as sweetened condensed milk, sugar, powdered margarine, and margarine. In addition, shop owners can also promote bundling cake boxes and cake mats. |
| 29 | @article{Fageeri2023MBAMB,  title={MBA: Market Basket Analysis Using Frequent Pattern Mining Techniques},  author={Sallam Osman Fageeri and Mohammad A. Kausar and Arockiasamy Soosaimanickam},  journal={International Journal on Recent and Innovation Trends in Computing and Communication},  year={2023},  url={https://api.semanticscholar.org/CorpusID:259884237}  } | **MBA: Market Basket Analysis Using Frequent Pattern Mining Techniques**  This Market Basket Analysis (MBA) is a data mining technique that uses frequent pattern mining algorithms to discover patterns of co-occurrence among items that are frequently purchased together. It is commonly used in retail and e-commerce businesses to generate association rules that describe the relationships between different items, and to make recommendations to customers based on their previous purchases. MBA is a powerful tool for identifying patterns of co-occurrence and generating insights that can improve sales and marketing strategies. Although a numerous works has been carried out to handle the computational cost for discovering the frequent itemsets, but it still needs more exploration and developments. In this paper, we introduce an efficient Bitwise-Based data structure technique for mining frequent pattern in large-scale databases. The algorithm scans the original database once, using the Bitwise-Based data representations as well as vertical database layout, compared to the well-known Apriori and FP-Growth algorithm. Bitwise-Based technique enhance the problems of multiple passes over the original database, hence, minimizes the execution time. Extensive experiments have been carried out to validate our technique, which outperform Apriori, Éclat, FP-growth, and H-mine in terms of execution time for Market Basket Analysis. |
| 30 | @article{Ghous2023DeepLB,  title={Deep Learning based Market Basket Analysis using Association Rules},  author={Hamid Ghous and Mubasher Malik and Iqra Rehman},  journal={KIET Journal of Computing and Information Sciences},  year={2023},  url={https://api.semanticscholar.org/CorpusID:264297776}  } | **Deep Learning based Market Basket Analysis using Association Rules**  Market Basket Analysis (MBA) is a data mining technique assisting retailers in determining the customer's buying habits while making new marketing decisions as the buyer's desire frequently changes with expanding needs; therefore, transactional data is getting large every day. There is a demand to implement Deep Learning (DL) methods to manipulate this rapidly growing data. In previous research, many authors conducted MBA applying DL and association rules (AR) on retail datasets. AR identifies the association between items to find in which order the customer place items in the basket. AR is only used in mining frequently purchased items from retail datasets. There is a gap in classifying these rules and predicting the next basket item using DL on the transactional dataset. This work proposes a framework using AR as a feature selection while applying DL methods for classification and prediction. The experiments were conducted on two datasets, InstaCart and real-life data from Bites Bakers, which operates as a growing store with three branches and 2233 products. The AR classified at 80,20 and 70,30 splits using CNNN, Bi- LSTM, and CNN-BiLSTM. The results considering simulation at both splits show that Bi-LSTM performs with high accuracy, around 0.92 on the InstaCart dataset. In contrast, CNN-BiLSTM performs best at an accuracy of around 0.77 on Bites Bakers dataset. |

**1.Giới thiệu bài toán**

**Đặt vấn đề**

Trong môi trường kinh doanh ngày nay, việc hiểu rõ hành vi mua hàng của khách hàng là yếu tố quyết định giữa thành công và thất bại của một doanh nghiệp bán lẻ. Phân tích giỏ hàng thị trường (Market Basket Analysis - MBA) là một công cụ quan trọng giúp các doanh nghiệp xác định các mẫu mua sắm và quy luật kết hợp từ dữ liệu giao dịch, từ đó tối ưu hóa các chiến lược kinh doanh.

Trong lĩnh vực phân tích giỏ hàng thị trường, thuật toán Apriori và FP-Growth đã trở thành hai công cụ phổ biến để khám phá các mẫu mua sắm. Tuy nhiên, việc lựa chọn thuật toán phù hợp và hiệu quả cho từng tình huống là một thách thức đối với các nhà kinh doanh và nhà nghiên cứu.

**Giới thiệu bài toán**

Trong bài nghiên cứu này, chúng ta sẽ tìm hiểu về phương pháp phân tích giỏ hàng thị trường thông qua hai thuật toán chính: Apriori và FP-Growth. Mục tiêu của chúng ta là so sánh hiệu suất và ứng dụng của hai thuật toán này trong việc khám phá mẫu mua sắm và quy luật kết hợp từ dữ liệu giao dịch của một doanh nghiệp bán lẻ.

Cụ thể, chúng ta sẽ:

* Trình bày cơ bản về các khái niệm, cách thức hoạt động và ứng dụng của phân tích giỏ hàng thị trường và hai thuật toán Apriori và FP-Growth trong quá trình này.
* Nghiên cứu và so sánh hiệu suất của hai thuật toán trong việc khám phá mẫu mua sắm từ một tập dữ liệu giao dịch thương mại điện tử.

Bằng cách tiếp cận này, chúng ta hy vọng có thể cung cấp cái nhìn sâu sắc hơn về sức mạnh và hạn chế của mỗi thuật toán, từ đó hỗ trợ các doanh nghiệp trong việc ra quyết định thông minh và hiệu quả cho các chiến lược kinh doanh của họ.

**2. Nguyên lý hoạt động**

MBA hoạt động dựa trên việc phân tích dữ liệu giao dịch bán hàng, bao gồm các thông tin như: sản phẩm được mua, số lượng sản phẩm được mua, giá trị của sản phẩm, thời gian mua hàng, thông tin khách hàng,... Từ dữ liệu giao dịch bán hàng, MBA có thể giúp xác định các mối quan hệ giữa các sản phẩm, như: các sản phẩm nào thường được mua cùng nhau, có khả năng thay thế cho nhau hoặc các bộ sản phẩm thường được mua vào một thời điểm nhất định. MBA có thể được sử dụng cho nhiều mục đích khác nhau như: cải thiện việc sắp xếp sản phẩm trong cửa hàng, phát triển các chương trình khuyến mãi hiệu quả, cá nhân hóa trải nghiệm mua sắm cho khách hàng và dự đoán xu hướng mua hàng trong tương lai. MBA là một công cụ mạnh mẽ có thể giúp các nhà bán lẻ tăng doanh thu, lợi nhuận và sự hài lòng của khách hàng.

Một số ví dụ về Phân tích giỏ thị trường:

Một cửa hàng tạp hóa có thể sử dụng MBA để xác định các sản phẩm thường được mua cùng nhau. Sau đó, cửa hàng có thể đặt các sản phẩm này cạnh nhau trên kệ để khuyến khích khách hàng mua thêm.

Một nhà bán lẻ quần áo có thể sử dụng MBA để xác định các sản phẩm có khả năng thay thế cho nhau. Sau đó, nhà bán lẻ có thể đề xuất các sản phẩm thay thế cho khách hàng khi họ hết hàng một sản phẩm nào đó.

Một trang web thương mại điện tử có thể sử dụng MBA để cá nhân hóa trải nghiệm mua sắm cho khách hàng. Trang web có thể đề xuất các sản phẩm phù hợp với sở thích mua sắm của khách hàng dựa trên lịch sử mua hàng của họ.

Có thể thấy, MBA là một công cụ có giá trị cho bất kỳ nhà bán lẻ nào muốn hiểu rõ hơn về hành vi mua hàng của khách hàng và đưa ra các quyết định kinh doanh hiệu quả hơn.

**3. Các nghiên cứu liên quan**

Phân tích giỏ hàng (Market Basket Analysis) là một phương pháp quan trọng trong lĩnh vực khoa học dữ liệu, cho phép các doanh nghiệp hiểu rõ hơn về hành vi mua sắm của khách hàng và tạo ra chiến lược kinh doanh hiệu quả. Đây là một công cụ quan trọng giúp cải thiện doanh số bán hàng bằng cách tối ưu hóa việc đặt hàng và chiến lược quảng cáo dựa trên thói quen tiêu dùng của khách hàng.

Trong phân tích giỏ hàng, thuật toán Apriori và FP-Growth là hai công cụ quan trọng được sử dụng để khám phá các mẫu mua sắm và quy luật kết hợp từ dữ liệu giao dịch. Apriori, một trong những thuật toán đầu tiên được giới thiệu, đã trở thành nền tảng trong phân tích giỏ hàng nhờ vào khả năng xác định các tập sản phẩm thường xuyên và mối quan hệ giữa chúng. Thông qua việc quét dữ liệu và kết hợp các mặt hàng một cách có hệ thống, Apriori tạo ra các tập sản phẩm lớn hơn và chỉ ra các xu hướng mua sắm phổ biến.

Tuy nhiên, Apriori có nhược điểm là tốn nhiều thời gian và dung lượng bộ nhớ do phải quét dữ liệu nhiều lần. Điều này làm hạn chế khả năng xử lý dữ liệu lớn hoặc thưa thớt. Để khắc phục vấn đề này, FP-Growth đã xuất hiện với ưu điểm vượt trội về hiệu suất và sử dụng ít bộ nhớ hơn. FP-Growth sử dụng cấu trúc dữ liệu FP-Tree để đại diện cho các mẫu mua sắm và tạo ra các tập sản phẩm thường xuyên một cách hiệu quả hơn. Mặc dù có nhược điểm khi xử lý dữ liệu thưa thớt, FP-Growth vẫn được ưa chuộng trong các tình huống đòi hỏi xử lý dữ liệu lớn nhanh chóng.

Kết hợp Apriori và FP-Growth cũng được nhiều nghiên cứu và ứng dụng thực tiễn chú ý. Việc kết hợp hai thuật toán này có thể tận dụng được ưu điểm của cả hai và giải quyết được nhược điểm riêng lẻ. Ví dụ, một nghiên cứu áp dụng cả hai thuật toán trong phân tích giỏ hàng cho thấy FP-Growth có thể xử lý dữ liệu nhanh hơn, trong khi Apriori đem lại kết quả chính xác hơn. Điều này làm rõ sự quan trọng của việc lựa chọn phương pháp phù hợp với mục tiêu nghiên cứu và tài nguyên có sẵn.

Tổng quan, phân tích giỏ hàng là một lĩnh vực quan trọng trong khoa học dữ liệu, và Apriori cùng FP-Growth là hai công cụ không thể thiếu trong quá trình này. Sự kết hợp giữa hai thuật toán này có thể mang lại nhiều lợi ích trong việc hiểu rõ hơn về hành vi mua sắm của khách hàng và tối ưu hóa chiến lược kinh doanh của doanh nghiệp.