|  |  |  |
| --- | --- | --- |
| ID | cite | Abstract |
| 1 | @article{Puka2021ANM,  title={A New Measure of Complementarity in Market Basket Data},  author={Radosław Puka and Stanislaw Jedrusik},  journal={J. Theor. Appl. Electron. Commer. Res.},  year={2021},  volume={16},  pages={670-681},  url={https://api.semanticscholar.org/CorpusID:234287806}  } | [A New Measure of Complementarity in Market Basket Data](https://www.semanticscholar.org/paper/A-New-Measure-of-Complementarity-in-Market-Basket-Puka-Jedrusik/479c44fd3a877f2eda24facec93b2031bed2dda1)  Modern IT systems collect detailed data on each activity, transaction, forum entry, conversation and many other areas. The availability of large data volumes in the business, industry and research fields opens up new opportunities for the empirical verification of various economic theories and laws. The analysis of big datasets in turn allows us to look at many issues from a new point of view and see the dependencies that are otherwise difficult to derive. In this paper, we propose a new measure for dependencies between goods in market basket data. The introduced measure was inspired by the well-known microeconomic concept of complementarity. Due to its similar properties to those of complementarity, the new measure was called basket complementarity (b-complementarity). B-complementarity not only measures the strength of dependencies between goods but also measures the direction of these dependencies. The values of the proposed measure can be relatively easily calculated using market basket data. This paper also presents a simple example illustrating this new concept, areas of possible application (e.g., in e-commerce) and preliminary results of searching for goods that meet the criteria of basket complementarity in real market basket data. |
| 2 | @article{Qisman2021MarketBA,  title={Market basket analysis using apriori algorithm to find consumer patterns in buying goods through transaction data (case study of Mizan computer retail stores)},  author={M Qisman and Rudi Rosadi and Atje Setiawan Abdullah},  journal={Journal of Physics: Conference Series},  year={2021},  volume={1722},  url={https://api.semanticscholar.org/CorpusID:234119984}  } | [Market basket analysis using apriori algorithm to find consumer patterns in buying goods through transaction data (case study of Mizan computer retail stores)](https://www.semanticscholar.org/paper/Market-basket-analysis-using-apriori-algorithm-to-Qisman-Rosadi/c1b90e983670e1bd0e410b4f159102c8ab1d5899)  Mizan Computer Shop is a shop that is engaged in the trading sector, especially in the field of selling computers and supporting accessories. Growing and increasing number of business actors in the computer sector, can makes the players challenged to be able to create unique differentiation and clear positioning. So, that consumers can differentiate from their competitors. Competitive and dynamic market conditions make every company should always observe competition in their business environment. Retail stores need to use all of available resources including data. Data processing is expected to be able to provide information that can be used to support marketing strategies. One of the data processing methods that are often used in marketing strategies is the use of data mining techniques i.e Market Basket Analysis using a priori algorithm. The application is designed using the waterfall method which starts from analyzing user needs, designing a process using UML which consists of: Use Case Diagrams, Activity Diagrams and Sequence Diagrams . This Market Basket Analysis application was built using the PHP programming language. From the results of the analysis in this research, it can be concluded that for the combination of 2 items with the highest confidence value 100% and the lift ratio value 3.39 i.e if a consumer buys a Laptop Charger, he will also buy a keyboard and for a combination of 3 items with the highest confidence value 100% and the lift ratio value 2.17 i.e if a consumer buys a Joystick and Laptop, he will also buy a mouse. |
| 3 | @article{Kraus2019PersonalizedPP,  title={Personalized Purchase Prediction of Market Baskets with Wasserstein-Based Sequence Matching},  author={Mathias Kraus and Stefan Feuerriegel},  journal={Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery \& Data Mining},  year={2019},  url={https://api.semanticscholar.org/CorpusID:170079161}  } | [Personalized Purchase Prediction of Market Baskets with Wasserstein-Based Sequence Matching](https://www.semanticscholar.org/paper/Personalized-Purchase-Prediction-of-Market-Baskets-Kraus-Feuerriegel/55135782847080a28eb877e16485ed6bd4c75840)  Personalization in marketing aims at improving the shopping experience of customers by tailoring services to individuals. In order to achieve this, businesses must be able to make personalized predictions regarding the next purchase. That is, one must forecast the exact list of items that will comprise the next purchase, \ie, the so-called market basket. Despite its relevance to firm operations, this problem has received surprisingly little attention in prior research, largely due to its inherent complexity. In fact, state-of-the-art approaches are limited to intuitive decision rules for pattern extraction, so that repeat purchases or co-purchases can be identified. However, the simplicity of the pre-coded rules impedes performance, since decision rules operate in an autoregressive fashion: the rules can only make inferences from past purchases of a single customer without taking into account the knowledge transfer that takes place between customers. In contrast, our research overcomes the limitations of pre-set rules by contributing a novel predictor of market baskets from sequential purchase histories: our predictions are based on similarity matching in order to identify similar purchase habits among the complete shopping histories of all customers. Our contributions are as follows: (1) We propose similarity matching based on subsequential dynamic time warping (SDTW) as a novel predictor of market baskets. Thereby, we can effectively identify cross-customer patterns. (2) We leverage the Wasserstein distance for measuring the similarity among embedded purchase histories. If desired, this can further be interpreted as a proxy to the prediction quality. (3) We develop a fast approximation algorithm for computing a lower bound of the Wasserstein distance in our setting. An extensive series of computational experiments demonstrates the effectiveness of our approach. The accuracy of identifying the exact market baskets based on state-of-the-art decision rules from the literature is outperformed by a factor of 4.0. This contributes to a further personalization in the provision of retail services. The actual use cases are widespread and include making customers tailored offerings in marketing, extending recommender systems for the purpose of suggesting personalized market baskets, and triggering product deliveries before purchase in order to accelerate delivery. |
| 4 | @inproceedings{Rusnandi2020PENERAPANDM, title={PENERAPAN KHAI THÁC DỮ LIỆU UNTUK ANALISIS THỊ TRƯỜNG GIỎ HÀNG DENGAN ALGORITMA FP-GROWTH PADA PD PASAR TOHAGA}, tác giả={Rusnandi Rusnandi và Suparni Suparni và Achmad Baroqah Pohan}, Year={2020}, url={https ://api.semanticscholar.org/CorpusID:226060066} } | [PENERAPAN DATA MINING UNTUK ANALISIS MARKET BASKET DENGAN ALGORITMA FP-GROWTH PADA PD PASAR TOHAGA](https://www.semanticscholar.org/paper/PENERAPAN-DATA-MINING-UNTUK-ANALISIS-MARKET-BASKET-Rusnandi-Suparni/f90e5c6a8ddae6b5b3f57e41a324c9071641a32e)  Sales data in 3 different shops (shop, Shop Maker Fernando and Son) at Tohaga Market in the form of PD book transactions are only seen in the absence of follow-up to determine the decision on who will come. Party owner only records the transactions of products sold and only see income per month. But with that data should be utilized to strategize on sales to come. By using the method of Frequent Pattern Growth Algorithm, the store can take decisions which require goods inventory more compared to other goods, and the placement of the goods in accordance with the relationship between the goods that are usually purchased a consumer can also be determined based on a Minimum Support and Minimum Confidence. Based on Market Basket Analysis obtained from the calculation of the Association by using the method of Frequent Pattern Growth Algorithm, then search for the value of the support and confidence to use Association Rules, Rules that are generated will be test by using Software RapidMiner. Then the placement of goods and inventory items in 3 different stores can be controlled with either the service so that the consumer will be increased, which in turn can increase the sales turnover. In this study Support is determined using threshold 40% and 83% Confidence. Having regard to the relationship of support and confidence the store owner can provide and put the items to be sold |
| 5 | @article{MAlawadh2022ASO,  title={A Survey on Methods and Applications of Intelligent Market Basket Analysis Based on Association Rule},  author={Monerah M. Alawadh and Ahmed M. Barnawi},  journal={Journal on Big Data},  year={2022},  url={https://api.semanticscholar.org/CorpusID:248651754}  } | [A Survey on Methods and Applications of Intelligent Market Basket Analysis Based on Association Rule](https://www.semanticscholar.org/paper/A-Survey-on-Methods-and-Applications-of-Intelligent-Alawadh-Barnawi/dca9df1892033eb1c6ef079aff62526e28a048fb)    The market trends rapidly changed over the last two decades. The primary reason is the newly created opportunities and the increased number of competitors competing to grasp market share using business analysis techniques. Market Basket Analysis has a tangible effect in facilitating current change in the market. Market Basket Analysis is one of the famous fields that deal with Big Data and Data Mining applications. MBA initially uses Association Rule Learning (ARL) as a mean for realization. ARL has a beneficial effect in providing a plenty benefit in analyzing the market data and understanding customers’ behavior. An important motive of using such techniques is maximizing the business profit as well as matching the exact customer needs as closely as possible. In this survey paper, we discussed several applications and methods of MBA based on ARL. Also, we reviewed some association rule learning measurements including trust, lift, leverage, and others. Furthermore, we discuss some open issues and future topics in the area of market basket analysis and association rule learning. |
| 6 | @article{Samboteng2022MARKETBA,  title={MARKET BASKET ANALYSIS OF ADMINISTRATIVE PATTERNS DATA OF CONSUMER PURCHASES USING DATA MINING TECHNOLOGY},  author={Lukman Samboteng and Rulinawaty Rulinawaty and M. Rachmat Kasmad and Mutmainnah Basit and Robbi Rahim},  journal={Journal of Applied Engineering Science},  year={2022},  url={https://api.semanticscholar.org/CorpusID:247142858}  } | [MARKET BASKET ANALYSIS OF ADMINISTRATIVE PATTERNS DATA OF CONSUMER PURCHASES USING DATA MINING TECHNOLOGY](https://www.semanticscholar.org/paper/MARKET-BASKET-ANALYSIS-OF-ADMINISTRATIVE-PATTERNS-Samboteng-Rulinawaty/c2649e0a06ad882c4f87544154d344940d0fb7ce)  Food is the ingredient that enables people to grow, develop, and achieve. For this reason, food quality and types of food must be considered so that they are safe for consumption and managed. Some plant-based foodstuffs are often processed and consumed by the community, even the most needed in food processing. In this case, the research was carried out using data mining with market basket analysis algorithms to obtain very valuable information to decide the inventory of the type of material needed. Market Based Analysis method is used to analyze all data and create patterns for each data. One method of Market Based Analysis in question is the association rule with a priori algorithm. This algorithm produces sales transactions with strong associations between items in the transaction which are used as sales recommendations that help users (owners) get recommendations when users see details of the itemset purchased. From the results of the trials in this study, it was found that the greater the minimum support (minsup) and minimum confidence (minconf), the less time it takes to produce recommendations and the fewer recommendations are given, but the recommendations given come from transactions that often appear. |
| 7 | @article{Chen2021ANM,  title={A New Method Combining Pattern Prediction and Preference Prediction for Next Basket Recommendation},  author={Guisheng Chen and Zhanshan Li},  journal={Entropy},  year={2021},  volume={23},  url={https://api.semanticscholar.org/CorpusID:242055380}  } | [A New Method Combining Pattern Prediction and Preference Prediction for Next Basket Recommendation](https://www.semanticscholar.org/paper/A-New-Method-Combining-Pattern-Prediction-and-for-Chen-Li/4e73232c79d88375a34a543547896d51f7dd99df)  Market basket prediction, which is the basis of product recommendation systems, is the concept of predicting what customers will buy in the next shopping basket based on analysis of their historical shopping records. Although product recommendation systems develop rapidly and have good performance in practice, state-of-the-art algorithms still have plenty of room for improvement. In this paper, we propose a new algorithm combining pattern prediction and preference prediction. In pattern prediction, sequential rules, periodic patterns and association rules are mined and probability models are established based on their statistical characteristics, e.g., the distribution of periods of a periodic pattern, to make a more precise prediction. Products that have a higher probability will have priority to be recommended. If the quantity of recommended products is insufficient, then we make a preference prediction to select more products. Preference prediction is based on the frequency and tendency of products that appear in customers’ individual shopping records, where tendency is a new concept to reflect the evolution of customers’ shopping preferences. Experiments show that our algorithm outperforms those of the baseline methods and state-of-the-art methods on three of four real-world transaction sequence datasets. |
| 8 | @article{Kurnia2019StudyOA,  title={Study of application of data mining market basket analysis for knowing sales pattern (association of items) at the O! Fish restaurant using apriori algorithm},  author={Yusuf Kurnia and Yohanes Isharianto and Yo Ceng Giap and Aditiya Hermawan and Riki},  journal={Journal of Physics: Conference Series},  year={2019},  volume={1175},  url={https://api.semanticscholar.org/CorpusID:209091449}  } | [Study of application of data mining market basket analysis for knowing sales pattern (association of items) at the O! Fish restaurant using apriori algorithm](https://www.semanticscholar.org/paper/Study-of-application-of-data-mining-market-basket-Kurnia-Isharianto/02e25417f32921a423a4c02dbda716aae8d0793f)  The development of the food and beverage culinary industry is growing very rapidly. Making food and beverage business owners, especially restaurants, have to make the right decision to stay in a very strong competition, restaurant owners must be ready to always innovate and remain to be able to meet consumer needs through products that can attract customers and determine strategies promotions that can boost sales. Stored transaction data has information that can be extracted by data mining techniques, for example knowing the pattern of sales in purchases by consumers. Information about sales patterns can be used by O! Fish restaurants to create more potential promotional strategies to boost sales by referring to items (menus) that are often purchased together. . To be able to find out the purchase patterns by consumers simultaneously, knowing what products are often purchased simultaneously can be used data mining techniques using a priori algorithms. A priori algorithm is used to generate association rules. Information about the association’s rules in purchasing items (menus) by consumers can be used by O! Fish restaurants to create more potential promotional strategies to boost sales by referring to a combination of items that are often purchased simultaneously. Later the results of this study are in the form of a website-based application to analyze purchasing patterns (item association rules) by consumers where the purchase pattern can be used as recommendations in determining the promotion development strategy for O! Fish restaurants. |
| 9 | @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 |
| 10 | @article{Wenninger2019DataMI,  title={Data Mining in Elite Beach Volleyball – Detecting Tactical Patterns Using Market Basket Analysis},  author={S. Wenninger and Daniel Link and Martin Lames},  journal={International Journal of Computer Science in Sport},  year={2019},  volume={18},  pages={1 - 19},  url={https://api.semanticscholar.org/CorpusID:202730345}  } | [Data Mining in Elite Beach Volleyball – Detecting Tactical Patterns Using Market Basket Analysis](https://www.semanticscholar.org/paper/Data-Mining-in-Elite-Beach-Volleyball-%E2%80%93-Detecting-Wenninger-Link/b79f55e1756012ae88be13d18d81a78437aefd90)  Sports coaches today have access to a growing amount of information that describes the performance of their players. Methods such as data mining have become increasingly useful tools to deal with the analytical demands of these high volumes of data. In this paper, we present a sports data mining approach using a combination of sequential association rule mining and clustering to extract useful information from a database of more than 400 high level beach volleyball games gathered at FIVB events in the years from 2013 to 2016 for both men and women. We regard each rally as a sequence of transactions including the tactical behaviours of the players. Use cases of our approach are shown by its application on the aggregated data for both genders and by analyzing the sequential patterns of a single player. Results indicate that sequential rule mining in conjunction with clustering can be a useful tool to reveal interesting patterns in beach volleyball performance data. |
| 11 | @article{Ojugo2019InventoryPA,  title={Inventory prediction and management in Nigeria using market basket analysis associative rule mining: memetic algorithm based approach},  author={A. Ojugo and A. Eboka},  journal={International Journal of Informatics and Communication Technology},  year={2019},  volume={8},  pages={128-138},  url={https://api.semanticscholar.org/CorpusID:211093801}  } | [Inventory prediction and management in Nigeria using market basket analysis associative rule mining: memetic algorithm based approach](https://www.semanticscholar.org/paper/Inventory-prediction-and-management-in-Nigeria-rule-Ojugo-Eboka/0dc71661ed4bbdaf4ce1514131fca5948f3d9ab5)  A key challenge in businesses today is determining inventory level for each product (to be) sold to clients. A pre-knowledge will suppress inventory stock-up and help avert unnecessary demurrage. It will also avoid stock out and avert loss of clients to competition. Study aims to unveil customer’s behavior in purchasing goods and thus, predict a next time purchase as well as serve as decision support to determine the required amount of each goods inventory. Study is conducted for Delta Mall (Asaba and Warri branches) department store. We adapt the memetic algorithm on market basket dataset to examine buying behavior of customers, their preference and frequency at which goods are purchased in common (basket). Result shows some items placed in basket allow customers to purchase items of similar value, or best combined with the selected items due to shelf-placement via concept of feature drift. Model yields 21-rules for eight items obtained from data transaction mining dataset acquired from Delta Mall. |
| 12 | @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](https://www.semanticscholar.org/paper/IMPLEMENTATION-OF-MARKET-BASKET-ANALYSIS-WITH-IN-Priyanto-Arifa/19a46fbff4e06bcec2d3492152995fb681f3e768)  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. |
| 13 | @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. |
| 14 | @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. |
| 15 | @article{Rana2021ASA,  title={A Seasonal and Multilevel Association Based Approach for Market Basket Analysis in Retail Supermarket},  author={S. Rana and Mohammad Nazrul Islam Mondal},  journal={European Journal of Information Technologies and Computer Science},  year={2021},  url={https://api.semanticscholar.org/CorpusID:244090920}  } | [A Seasonal and Multilevel Association Based Approach for Market Basket Analysis in Retail Supermarket](https://www.semanticscholar.org/paper/A-Seasonal-and-Multilevel-Association-Based-for-in-Rana-Mondal/58542099d41dfc47c6d2a1a4efac44faa52d704a)  Market Basket Analysis is an observational data mining methodology to investigate the consumer buying behavior patterns in retail Supermarket. It analyzes customer baskets and explores the relationship among products that helps retailers to design store layouts, make various strategic plans and other merchandising decisions that have a big impact on retail marketing and sales. Frequent itemsets mining is the first step for market basket analysis. The association rules mining uncovers the relationship among products by looking what products the customers frequently purchase together. In retail marketing, the transactional database consists of many itemsets that are frequent only in a particular season however not taken into consideration as frequent in general. In some cases, association rules mining at lower data level with uniform support doesn't reflect any significant pattern however there is valuable information hiding behind it. To overcome those problems, we propose a methodology for mining seasonally frequent patterns and association rules with multilevel data environments. Our main contribution is to discover the hidden seasonal itemsets and extract the seasonal associations among products in additionally with the traditional strong regular rules in transactional database that shows the superiority for making season based merchandising decisions. The dataset has been generated from the transaction slips in large supermarket of Bangladesh that discover 442 more seasonal patterns as well as 1032 seasonal association rules in additionally with the regular rules for 0.1% minimum support and 50% minimum confidence. |
| 16 | @inproceedings{Kabasakal2020UnderstandingSB,  title={Understanding Shopping Behaviors With Category- and Brand-Level Market Basket Analysis},  author={İnanç Kabasakal},  year={2020},  url={https://api.semanticscholar.org/CorpusID:202354523}  } | [Understanding Shopping Behaviors With Category- and Brand-Level Market Basket Analysis](https://www.semanticscholar.org/paper/Understanding-Shopping-Behaviors-With-Category-and-Kabasakal/27d131566eaea3637998c6e577c945eff98d8d10)  In a digital transformation environment, most businesses shift towards e-business and encounter businesses and customer interaction on digital channels. Information Technology renders data access and processing more efficient, and use of customer data in decision making has become a focal interest area that attracts researchers. Customer data is a relevant subject for numerous studies in Data Mining. In this chapter, Association Rule Mining has been utilized to extract purchase behavior patterns with a multilevel approach. Basket data obtained from an online retailer was analyzed to discover purchase behaviors with a focus on category and brand attributes of products. Brands and categories purchased together frequently were discovered. Brand and category-wise association rules were also presented in the results. The analysis differs from the majority of prior analyses, by referring to the category and brand attributes in basket data. It could be noted that generalized rules obtained with this approach might prove useful in recommending new items of existing brands or categories. |
| 17 | @article{Sarada2019RealizingBP,  title={Realizing Behavioral Patterns using Fuzzy Logic in Market Basket Analysis},  author={Weragoda Sarada and P. V. Kumar},  journal={International journal of engineering research and technology},  year={2019},  volume={8},  url={https://api.semanticscholar.org/CorpusID:212825021}  } | [Realizing Behavioral Patterns using Fuzzy Logic in Market Basket Analysis](https://www.semanticscholar.org/paper/Realizing-Behavioral-Patterns-using-Fuzzy-Logic-in-Sarada-Kumar/7e77694e53af2755a06dc5a89743d042f1ef0056)  Data mining is an area of research and study within a computer science discipline involving to make out the meaning and interpret the information or data, something that repeats in a predictable way which refer to a design or to customary behavior through any type of calculation that includes both arithmetical and non-arithmetical steps and follows a well-defined model, for example an algorithm, that integrates technology with a plan or intention or an idea or invention to help sell or publicize a commodity in view of such as AI, database systems, ML and statistics. In this paper, the focus is on a new novel approach which scans the market basket database and finds those occurrences of items which are of noteworthy and prunes consecutive item sets which satisfy the support and confidence threshold, interest and generates association rules which are used to locate the way in which two or more things are connected in the vast database that realizes the buy conduct of the client and helps in the expanding of the deals in the grocer’s stores or super market and also in choosing the precise territory and the accurate period of gathering in crops, farming and helps in increasing the produce. |
| 18 | @inproceedings{2019MarketBA,  title={Market Basket Analysis with Enhanced Support Vector Machine (ESVM) Classifier for Key Security in Organization},  author={},  year={2019},  url={https://api.semanticscholar.org/CorpusID:219555707}  } | [Market Basket Analysis with Enhanced Support Vector Machine (ESVM) Classifier for Key Security in Organization](https://www.semanticscholar.org/paper/Market-Basket-Analysis-with-Enhanced-Support-Vector/a7135142d724f22e30be5ea8201c36edaa7f08ac)  Market Basket Analysis is considered to be one among the highly popular and efficient sort of data analysis exploited in the marketing and retailing field. The objective of market basket analysis lies in deciding the products purchased together by the customers. Its name has originated from the concept of customers filling into a shopping cart everything of all they had purchased (a "market basket") while doing shopping in the grocery. Having a knowledge of the products that customers buy in group can be quiteusefulfor a retailer or to any other organization. A store could make the best use of this information to keep the products that are often sold together in the same place, whereas a catalog or World Wide Web (WWW) merchant could utilize it for deciding the structure of their catalog and order form. Since several applications such as market basket analysis, fraud detection in web, medical diagnosis, census data, Customer Relationship Management of business that makes use of association rules exists, the process involving Decision making can be improved. Security is also regarded to bean important facet for transactions done individually and frequent itemsets for database that are horizontally partitioned. In order to render security for lastly bough often used itemsets for transaction purposes, this research work introduces a novel key security algorithm that uses RSA cryptographic technique which is classifier based. The classifier makes use of information about several often utilized itemsets and it provides a key value to the actual company. For instance, in case if there are any reliance users, only the valid users can obtain that market info. The rest of the users belonging to the reliance organization are not allowed to select the data’s key value. First, the frequent itemsets are mined with the help of association rule mining employing Probabilistic Graphical Model techniques. Then the Enhanced Support Vector Machine (ESVM) classifier checks the key values of the mined frequent itemsets. |
| 19 | @inproceedings{Wang2019MarketBA,  title={Market Basket Analysis based on Apriori and CART},  author={Liyuan Wang and Jian Qing Sun},  year={2019},  url={https://api.semanticscholar.org/CorpusID:198946351}  } | [Market Basket Analysis based on Apriori and CART](https://www.semanticscholar.org/paper/Market-Basket-Analysis-based-on-Apriori-and-CART-Wang-Sun/8c2648cdb05fd202001e262aa5f309bd1f416315)  With the rapid development of economy and information technology, the development of the market retail industry cannot be underestimated. How can it improve the efficiency of the retail industry? The paper which uses Apriori algorithm to find out the data of shopping basket from the massive data of consumers reveals the relationship between the purchased goods, and subsequently applies the association rules and CART decision tree algorithm to reveal the characteristics of the customer group and the target customers classification. In order to dig out more detailed and valuable information, it is convenient for the goods to be better configured and sold, and to improve the operational efficiency of the market. |
| 20 | @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. |
| 21 | @inproceedings{Javier2022TemporalVP,  title={Temporal Visual Profiling of Market Basket Analysis},  author={Francisco Javier and Moreno Arboleda and Giovanni P{\'e}rez Ortega and Jaime Alberto and Guzm{\'a}n Luna},  year={2022},  url={https://api.semanticscholar.org/CorpusID:260515634}  } | [Temporal Visual Profiling of Market Basket Analysis](https://www.semanticscholar.org/paper/Temporal-Visual-Profiling-of-Market-Basket-Analysis-Javier-Arboleda/3b4515d936cacebe7967d1b48741f0ad0061ca20)  Market basket analysis allows analysts to understand the behaviour of customers. In this paper, we propose a novel technique to generate the profile of a customer with regard to his/her product purchase history. After obtaining the profile of a customer, we present a visual technique, to compare in a friendly and interactive way the profiles of customers over time. To show the expediency of our proposal, we performed experiments with a dataset of purchases of a retail market. The results showed that our proposal can be useful for stock planning and identifying customers with similar buying tendencies. |
| 22 | @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. |
| 23 | @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{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](https://www.semanticscholar.org/paper/E-Commerce-Market-Basket-Analysis-using-Apriori-Gupta-Shah/eac521d3dd4d7f51d23510a9392be0d7c45b5a0b)  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 |
| 25 | @article{Umar2022MarketBA,  title={Market Basket Analysis Menggunakan Association Rule dan Algoritma Apriori Pada Produk Penjualan Mitra Swalayan Salatiga},  author={Elfira Umar and Daniel H. F. Manongga and Ade Iriani},  journal={JURNAL MEDIA INFORMATIKA BUDIDARMA},  year={2022},  url={https://api.semanticscholar.org/CorpusID:253001514}  } | [Market Basket Analysis Menggunakan Association Rule dan Algoritma Apriori Pada Produk Penjualan Mitra Swalayan Salatiga](https://www.semanticscholar.org/paper/Market-Basket-Analysis-Menggunakan-Association-Rule-Umar-Manongga/47255612f9d43b79336ecfe1046f2aebfce66bdc)  Market Basket analysis is learning to manage associations in data processing in various fields. The main purpose of Market Basket analysis in the field of sales is to convey an important message to the company so that it can find out the behavior patterns of entering goods into the shopping basket by consumers so that partners can make a decision. In this study, the Apriori Algorithm is used to take into account changes that occur in the data. This study discusses data mining techniques in analyzing what items are most often purchased at the same time by consumers so that they can change the placement of items that are close together to increase the impulse buying effect. The results obtained are 5 rules where one of the rules obtains the highest confidence value when buying cigarettes, the dominant item is taken simultaneously, namely eggs by obtaining a confidence value that can meet the highest confidence requirements, namely 67%. |
| 26 | @article{Rao2021ApplicationOM,  title={Application of market–basket analysis on healthcare},  author={Abishek B. Rao and Jammula Surya Kiran and Poornalatha G},  journal={International Journal of System Assurance Engineering and Management},  year={2021},  volume={14},  pages={924-929},  url={https://api.semanticscholar.org/CorpusID:239664123}  } | [Application of market–basket analysis on healthcare](https://www.semanticscholar.org/paper/Application-of-market%E2%80%93basket-analysis-on-healthcare-Rao-Kiran/d47b3f97f4514bf54b3e0087012cd9fc14b9be38)  Data analysis plays a vital role in the present era as it helps us to understand the patterns by exploring it in meaningful ways. Market—basket is one of the main methods used to find frequently occurring items in a transactional database and many researchers use the Apriori algorithm for this purpose. This paper presents the application of Market Basket Analysis to the healthcare section. The present work tries to find frequent diseases that occur together in an area by using the Apriori algorithm. This could help the residents of an area to be more cautious about the frequently occurring diseases and take all possible precautionary measures to safeguard their health. In addition, it could also help the doctors so that, they are ready with required medications to treat the patients. |
| 27 | @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. |
| 28 | @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](https://www.semanticscholar.org/paper/BASKET-MARKET-ANALYSIS-USING-R-BASED-APRIORI-TO-Yustiana-Sutarha/294a1df96a4e52f747136895dea983353e5b568c)  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 |
| 29 | @inproceedings{Prawira2020MarketBA,  title={Market Basket Analysis To Identify Stock Handling Patterns \& Item Arrangement Patterns Using Apriori Algorithms},  author={Tresna Yudha Prawira and Sunardi Sunardi and Abdul Fadlil},  year={2020},  url={https://api.semanticscholar.org/CorpusID:219102084}  } | [Market Basket Analysis To Identify Stock Handling Patterns & Item Arrangement Patterns Using Apriori Algorithms](https://www.semanticscholar.org/paper/Market-Basket-Analysis-To-Identify-Stock-Handling-%26-Prawira-Sunardi/3fcc86958582c2c1f54259028e6a8f38c4d194fd)  The process of managing the pattern of handling stock of goods and the pattern of arranging goods on store shelves requires an identification process by utilizing data from sales transaction results. Market basket analysis of sales transaction data using Apriori Algorithm stages produces an information in the form of association rules with a minimum support value of 50% and a minimum confidence of 60%. It can be a reference in the arrangement of items on store shelves by referring to a combination of items that are often bought by consumers simultaneously. In addition, the stock inventory pattern can take advantage of the results of determining the high frequency value in the combination pattern 1 - itemset C1 with a minimum support value of 50% which is compared with the initial inventory. |
| 30 | @article{Wahyudi2023MarketBA,  title={Market Basket Analysis dengan Perbandingan Metode Apriori dan FP-Growth Pada Data Transaksi XYZ},  author={Rizki Nofrian Wahyudi and Dyah Erny Herwindiati and Janson Hendryli},  journal={Jurnal Ilmu Komputer dan Sistem Informasi},  year={2023},  url={https://api.semanticscholar.org/CorpusID:260186039}  } | [Market Basket Analysis dengan Perbandingan Metode Apriori dan FP-Growth Pada Data Transaksi XYZ](https://www.semanticscholar.org/paper/Market-Basket-Analysis-dengan-Perbandingan-Metode-Wahyudi-Herwindiati/bd51667740b5d6ffa3e08d7beffaa3799112f00f)  Technology is currently advancing quickly, allowing all organizations to grow their networks with its aid and create sales methods that now rely on technology to aid in making the proper judgments. When saved transaction data is accessible, every business will be able to implement its marketing strategy to maximize client transactions. use it to your advantage. Analysis of the market basket using the FP-Growth and a priori algorithm in transactions that aid in strategic planning and business product structuring. The FP-Growth algorithm and the Apriori algorithm work well together. One can evaluate the effectiveness of the employment of the a priori algorithm and the FP-Growth algorithm by applying both of them |
| 31 | @article{Firmansyah2021MarketBA,  title={Market Basket Analysis for Books Sales Promotion using FP Growth Algorithm, Case Study : Gramedia Matraman Jakarta},  author={Firmansyah Firmansyah and Agus Yulianto},  journal={Journal of Information Technology Education},  year={2021},  volume={4},  pages={383-392},  url={https://api.semanticscholar.org/CorpusID:234151819}  } | [Market Basket Analysis for Books Sales Promotion using FP Growth Algorithm, Case Study : Gramedia Matraman Jakarta](https://www.semanticscholar.org/paper/Market-Basket-Analysis-for-Books-Sales-Promotion-FP-Firmansyah-Yulianto/2019ccddf09331b69b573fd891689c309d1d2a1e)  For retail companies such as Gramedia stores, promotion and strategies to sell books are important, so tools are needed to analyze past sales data. Gramedia does not yet have tools to analyze shopping cart patterns that aim to carry out product promotions appropriately. To promote what books should be promoted using the market basket analysis method or shopping basket analysis. The algorithm used in the data mining process is Frequent Pattern Growth (FP Growth) because it is faster in processing large data. The data analyzed is historical data on book sales from January to March 2020 which is taken randomly (random sampling). The framework used in the data mining process is the Cross Industry Standard Process for Data Mining (CRISP-DM) and the tool used is the Rapid Miner using a market basket analysis framework. With a minimum support of 0.003 and a minimum confidence 0.3 using the FP-Growth algorithm to produce an item set of 7 rules to recommend product promotions. The algorithm results are also in accordance with the business understanding phase of CRISP-DM. |
| 32 | @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. |
| 33 | @article{HarshaliPatil2023EnhancingRS,  title={Enhancing Retail Strategies through Apriori, ECLAT\& FP Growth Algorithms in Market Basket Analysis},  author={Et al. Harshali Patil,},  journal={International Journal on Recent and Innovation Trends in Computing and Communication},  year={2023},  url={https://api.semanticscholar.org/CorpusID:267330786}  } | [Enhancing Retail Strategies through Apriori, ECLAT& FP Growth Algorithms in Market Basket Analysis](https://www.semanticscholar.org/paper/Enhancing-Retail-Strategies-through-Apriori%2C-ECLAT%26-Et-al.-Harshali-Patil/53226f6882ed102e45a728072f78334f077614f8)  "Market basket analysis" is a method employed in data mining to discover items that are commonly bought together by customers in a retail store. It is a crucial tool for retailers to understand consumer purchasing behavior and to improve their sales and marketing strategies. In this research paper, we present a comprehensive study on market basket analysis using three popular algorithms: Apriori, ECLAT, and FPGrowth. The paper begins with a brief synopsis of market basket analysis and the techniques adopted for itemset mining. We then introduce the dataset used in this study, which consists of real-life transaction data collected from a retail store. Next, we perform a thorough evaluation of the Apriori, ECLAT, and FPGrowth algorithms in terms of their computational time and the quality of the association rules generated. The results show that the FPGrowth algorithm is the fastest of the three algorithms, while the Apriori algorithm generates the most comprehensive and high-quality association rules.In addition, we also present a comparison of the performance of these algorithms that involve different assessment criteria like support, confidence, and lift. Our study highlights the importance of selecting the appropriate algorithm for market basket analysis depending on the specific requirements and constraints of the task. The paper concludes with an analysis on the limitations and future directions of research in this area. Overall, our study provides insights into the strengths and weaknesses of the Apriori, ECLAT, and FPGrowth algorithms and functions as a valuable resource for professionals and researchers in the field of market basket analysis. |
| 34 | @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. |
| 35 | @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](https://www.semanticscholar.org/paper/Comparison-of-Apriori%2C-Apriori-TID-and-FP-Growth-in-Idris-Sampetoding/8bb6e760081abf257b9f9f9c82bdcc0a63e56803)  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 |

**PHÂN TÍCH GIỎ HÀNG ĐỂ NGHIÊN CỨU HÀNH VI KHÁCH HÀNG**

**I. GIỚI THIỆU**

Thời đại số mang đến cho các nhà bán lẻ vô số cơ hội mới để thu thập dữ liệu và phân tích giỏ thị trường. Với sự bùng nổ của thương mại điện tử, các nhà bán lẻ có thể thu thập dữ liệu về hành vi mua hàng của khách hàng một cách chi tiết và chính xác hơn bao giờ hết. Việc tìm hiểu và dự đoán quá trình ra quyết định của người tiêu dùng đã và đang được nghiên cứu rộng rãi. Quá trình ra quyết định của họ dẫn đến việc hình thành giỏ hàng đa danh mục, bao gồm tất cả các mặt hàng mà một cá nhân mua cùng lúc. Giỏ hàng thị trường này là mục tiêu quan tâm của cả nhà bán lẻ trực tuyến và trực tiếp, bởi nó cung cấp thông tin cho doanh nghiệp để triển khai các chương trình tiếp thị cá nhân hóa và bán chéo hiệu quả. Nhu cầu này càng được thúc đẩy bởi xu hướng cá nhân hóa trong marketing hiện nay, điều đó dẫn đến sự ra đời của nhiều ứng dụng dự đoán [3]. Dữ liệu thu thập được từ các kênh trực tuyến như trang web bán hàng, ứng dụng di động và mạng xã hội có thể được sử dụng để phân tích giỏ thị trường theo nhiều cách khác nhau. Phân tích giỏ hàng là một tập hợp các kỹ thuật phân tích nhằm khám phá mối quan hệ và tương quan giữa các sản phẩm thông qua việc phân tích giỏ hàng của khách hàng [24]. Việc phân tích các tập dữ liệu lớn cho phép chúng ta nhìn nhận nhiều vấn đề từ góc độ mới và thấy được những mối quan hệ giữa các sản phẩm vốn khó phát hiện. Trong lĩnh vực kinh doanh, kỹ thuật này được sử dụng để điều tra xem sản phẩm nào khác có khả năng được mua kèm cao hơn khi khách hàng mua một sản phẩm nhất định. Ví dụ, khi khách hàng mua mì gói, thì trong hầu hết các trường hợp, họ cũng sẽ mua trứng. Việc khai thác các quy tắc liên kết như vậy giữa các sản phẩm trong thị trường bán lẻ đóng vai trò quan trọng trong việc: xây dựng các chiến lược kinh doanh mới, thiết kế các chương trình khuyến mãi, giảm giá hoặc thiết kế bố cục cửa hàng. Phân tích giỏ thị trường (MBA) là một kỹ thuật khai thác dữ liệu mạnh mẽ được sử dụng trong lĩnh vực bán lẻ để hiểu các mô hình hành vi mua hàng của khách hàng. Nó phân tích dữ liệu giao dịch, thường được ghi lại tại điểm bán hàng, bằng cách sử dụng phương pháp Thuật toán tăng trưởng mẫu thường xuyên, cửa hàng có thể đưa ra các quyết định yêu cầu tồn kho hàng hóa nhiều hơn so với các hàng hóa khác và việc sắp xếp hàng hóa phù hợp với mối quan hệ giữa các hàng hóa thường được người tiêu dùng mua cũng có thể được xác định dựa trên về sự hỗ trợ tối thiểu và sự tin cậy tối thiểu[4]. MBA có thể giúp doanh nghiệp đo lường mức bổ sung trong dữ liệu giỏ hàng**,** đề xuất một thước đo mới để đánh giá mối quan hệ phụ thuộc giữa các mặt hàng trong dữ liệu giỏ hàng, giúp xác định chính xác hơn những sản phẩm thường được mua cùng nhau [1]. Mô hình [25] nhằm mục đích dự đoán việc bổ sung và không bổ sung các mặt hàng vào giỏ hàng thị trường của khách hàng, nhằm giúp các nhà tiếp thị hành động thuận tiện, ví dụ như đề xuất các mặt hàng phù hợp hơn. Nó nhằm mục đích giúp các nhà quản lý cải thiện mối quan hệ với khách hàng đồng thời tăng doanh số bán hàng[9].Trong những năm gần đây, các nhà phân tích đã áp dụng nhiều thuật toán trong MBA để xác định các sản phẩm thường được mua cùng nhau, hỗ trợ xây dựng chiến lược khuyến mãi hiệu quả. Kết quả là, các công ty đặt mục tiêu dự đoán giỏ hàng cá nhân hóa từ lần mua hàng tiếp theo của cá nhân, dựa trên đó họ cải thiện dịch vụ khách hàng, quản lý chuỗi cung ứng hoặc tối ưu hóa phân loại sản phẩm. Phân tích giỏ thị trường không chỉ đơn thuần là một thuật toán tìm kiếm các mặt hàng thường được mua cùng nhau mà còn là một công cụ phân tích mạnh mẽ, giúp các doanh nghiệp thấu hiểu hành vi khách hàng, đưa ra những chiến lược marketing phù hợp, từ đó gia tăng doanh số và lợi nhuận.

**II. NGHIÊN CỨU LIÊN QUAN**

Phân tích giỏ hàng (Market Basket Analysis) là kĩ thuật phân tích hành vi khách hàng dựa trên lịch sử giao dịch của người mua hàng, từ đó giúp cho bộ phận kinh doanh nắm được thị hiếu, thói quen tiêu dùng của khách hàng để có những chiến dịch Marketing một cách hợp lý. Thuật toán này được sử dụng để cải thiện doanh số của các cửa hàng bán nhiều sản phẩm bằng cách sắp xếp vị trí sản phẩm dựa trên thói quen tiêu dùng của khách hàng [24]. Kỹ thuật hoạt động dựa trên việc khám phá các mối tương quan giữa các sản phẩm thường được mua cùng nhau trong giỏ hàng của người tiêu dùng. Hiện nay, Phân tích Giỏ Hàng ngày càng được ứng dụng rộng rãi, nó là nền tảng cho nhiều kỹ thuật và thuật toán khai thác dữ liệu. Nhiều thuật toán khai thác dữ liệu được sử dụng trong Phân tích Giỏ Hàng, trong đó Apriori, AIS, SETM và FP-Growth là những ví dụ nổi bật. Mặc dù các kỹ thuật mới liên tục xuất hiện, Apriori và FP-Growth vẫn là hai thuật toán phổ biến nhất do hiệu quả và tính ứng dụng cao.

Thuật toán Apriori là công cụ được sử dụng rộng rãi trong khai thác luật kết hợp, đặc biệt là trong Phân tích Giỏ hàng. Các thuật toán như AI và SETM cho rằng Apriori có độ chính xác cao. Apriori xác định các tập sản phẩm thường xuyên (frequent itemsets) trong giao dịch và xác định mối quan hệ (association rules) giữa các sản phẩm này. Thuật toán được áp dụng trong nhiều nghiên cứu về những lĩnh vực khác nhau, như: Phân tích Giỏ Hàng trên nền tảng Thương mại Điện tử [30], Phân tích Giỏ Hàng tại Cửa hàng Tiện lợi [27], Phân tích Giỏ Hàng trong Y tế [28], Phân tích Giỏ Hàng tại Cửa hàng Máy tính [29],… Thuật toán Apriori là nền tảng trong các hệ thống cơ sở dữ liệu quan hệ, được sử dụng chủ yếu để khám phá các mẫu sản phẩm thường được mua cùng nhau thông qua quy tắc kết hợp (association rule learning) [26]. Mục tiêu chính của thuật toán là xác định các mặt hàng xuất hiện với tần suất lớn trong cơ sở dữ liệu, sau đó kết hợp tăng dần để hình thành các tập sản phẩm lớn hơn, với điều kiện các sự kết hợp này xuất hiện *nhất quán trong dữ liệu*. Thuật toán Apriori có vai trò quan trọng trong việc Phân tích Giỏ hàng vì các tập sản phẩm thường xuyên được thuật toán Apriori xác định chính là nền tảng cho việc tạo ra các luật kết hợp, từ đó chỉ ra các xu hướng chi phối trong cơ sở dữ liệu. Thuật toán Apriori xem xét mỗi giao dịch như một tập sản phẩm riêng biệt. Kỹ thuật này sử dụng “phương pháp từ dưới lên” (bottom-up approach), bắt đầu với các sản phẩm xuất hiện thường xuyên riêng lẻ, rồi kết hợp chúng một cách có hệ thống bằng cách tham chiếu thông tin từ dữ liệu. Quá trình kết hợp dừng khi không tìm được sự kết hợp mới nào có ý nghĩa [26].

Thuật toán Apriori được biết đến là thuật toán lâu đời và đơn giản, tuy nhiên lại tốn nhiều thời gian tính toán và dung lượng bộ nhớ để tìm kiếm các tập sản phẩm (itemsets) [30]. Điều này là do thuật toán phải quét dữ liệu lặp đi lặp lại nhiều lần. Thuật toán này được sử dụng để xác định mối tương quan giữa các mặt hàng mà khách hàng quan tâm, lưu trữ trong cơ sở dữ liệu. Sau khi thu được các tập sản phẩm được mua thường xuyên, một quy tắc được chọn và sau đó các nhà nghiên cứu phân tích, so sánh về tốc độ thực thi, quá trình tạo quy tắc và độ chính xác của quy tắc từ thuật toán [30]. Phân tích giỏ hàng của dữ liệu giao dịch bán hàng sử dụng các giai đoạn của Thuật toán Apriori sẽ tạo ra thông tin dưới dạng các quy tắc kết hợp với giá trị hỗ trợ tối thiểu là 50% và độ tin cậy tối thiểu là 60% [31]. Đây có thể là tài liệu tham khảo trong việc sắp xếp các mặt hàng trên kệ cửa hàng bằng cách tham khảo sự kết hợp các mặt hàng thường được người tiêu dùng mua đồng thời. Thuật toán Apriori là một công cụ hữu ích để khai thác tri thức từ dữ liệu, đặc biệt là trong lĩnh vực phân tích giỏ hàng (market basket analysis). Tuy nhiên, thuật toán có nhược điểm là tốn nhiều thời gian và dung lượng, do đó bị hạn chế sử dụng cho các tập dữ liệu lớn hoặc thưa thớt.

Thuật toán Thuật toán FP-Growth (Frequent Pattern Growth) được nhiều nhà phân tích chọn lựa vì khắc phục được khuyết điểm của Thuật toán Apriori [33]. Thuật toán này có 3 bước chính: Xây dựng cây FP (FP-Tree Construction), Phát triển FP-Growth (FP-Growth Generation), Tập mục thường xuyên (Frequent Item Set) [33]. So với Apriori, FP-Growth có nhiều ưu điểm nổi trội như hiệu quả xử lý dữ liệu lớn nhanh hơn và sử dụng ít bộ nhớ hơn. Tuy nhiên, thuật toán này cũng có một số nhược điểm nhất định như gặp khó khăn trong việc xử lý dữ liệu thưa thớt và yêu cầu nhiều thời gian để xây dựng cây FP-Tree [32].

Ngoài việc áp dụng đơn lẻ hai thuật toán trên, nhiều nhà nghiên cứu còn ứng dụng cả 2 thuật toán vào việc Phân tích Giỏ hàng. Một công ty dược phẩm ở Indonesia đã áp dụng các kỹ thuật khai thác dữ liệu, đặc biệt là kết hợp sử dụng thuật toán Apriori và FP-GROWTH cho phân tích giỏ hàng trên PT. XYZ [36]. Kết quả cho thấy: FP-GROWTH có thời gian thực hiện tác vụ nhanh nhất với 84.655 giây. Tuy nhiên, việc sử dụng bộ nhớ cho thuật toán Apriori là thấp nhất ở mức 482,32 MiB, với mức tăng 0,21 MiB. Đối với các quy tắc được tạo, cả hai thuật toán đều tạo ra số lượng quy tắc, giá trị hỗ trợ, độ tin cậy, độ nâng, độ hỗ trợ Bi, độ tin cậy Bi và độ nâng Bi tương tự nhau. Từ đó, các nhà nghiên cứu kết luận: với các tập dữ liệu yêu cầu tính dễ triển khai và sử dụng bộ nhớ thì nên sử dụng thuật toán Apriori. Nếu tốc độ thực thi và khối lượng dữ liệu lớn được ưu tiên thì FP-GROWTH sẽ là lựa chọn tốt hơn vì thời gian thực thi nhanh hơn cho khối lượng dữ liệu lớn [36].

Nghiên cứu này khảo sát vai trò của khoa học dữ liệu trong việc tìm hiểu nhu cầu khách hàng và nâng cao doanh số, tập trung cụ thể vào việc áp dụng thuật toán Apriori và FP-Growth tại cửa hàng bán lẻ Deli Point ở Labuan Bajo. Nghiên cứu làm sáng tỏ tác động của "dữ liệu rác" đối với việc phân tích dữ liệu giao dịch, nhấn mạnh sự cần thiết của các quy trình làm sạch dữ liệu nhằm đảm bảo kết quả chính xác. Nghiên cứu chỉ ra thuật toán FP-Growth có lợi thế về tốc độ xử lý trong khi thuật toán Apriori đưa ra kết quả chính xác. Nghiên cứu này đóng góp những hiểu biết có giá trị về hành vi của người tiêu dùng bán lẻ và các chiến lược đặt sản phẩm hiệu quả, đồng thời nhấn mạnh vai trò chuyển đổi của khoa học dữ liệu trong việc tối ưu hóa doanh số và thúc đẩy khả năng cạnh tranh trong lĩnh vực bán lẻ. Nhìn chung, việc kết hợp Apriori và FP-Growth có thể mang lại nhiều lợi ích cho phân tích giỏ hàng, nhưng cũng cần cân nhắc đến những hạn chế đi kèm. Cần lựa chọn phương pháp phù hợp với mục tiêu nghiên cứu, kích thước dữ liệu và tài nguyên sẵn có.