Issues in Information Systems
Volume 25, Issue 2, pp. 363-378, 2024

DOI: https://doi.org/10.48009/2_iis_2024_129

Leveraging market basket analysis for enhanced understanding of social media platform usage

Alan Peslak, *Penn State University, arp14@psu.edu*
Pratibha Menon, *Pennsylvania Western University, menon@pennwest.edu*

Abstract

This paper explores the application of Market Basket Analysis (MBA), a quantitative technique traditionally employed in retail environments to enhance marketing and inventory strategies, to the realm of social media usage. Recognizing the potential of MBA to reveal insightful patterns within complex datasets, this study adapts the approach to analyze interactions across diverse social media platforms, thus framing user activities on these platforms as analogous to consumer transactions in a market basket. Leveraging a dataset provided by Pew Research, which encompasses responses from adults across all 50 U.S. states, the study employs the Apriori algorithm to identify and evaluate associations between major platforms such as Facebook, Instagram, and Twitter. The analysis focuses on deriving metrics such as support, confidence, lift, and deployability to ascertain the robustness and relevance of the identified associations. The results underscore significant interconnections between user engagements across different platforms, providing empirical support for enhanced strategic decision-making by platform developers and marketers. By extending the application of MBA to the study of digital interactions, this paper not only broadens the understanding of user behavior in digital ecosystems but also enriches the methodological toolkit available for digital sociology and media studies, demonstrating the adaptability of MBA across varied analytical contexts.

Keywords: market-basket-analysis, social-media-platform, support, confidence, lift, deployability.

Introduction

In the United States, at least 72% of adults use various social media platforms (Pew Research Center, 2021). Popular social networking sites and applications include Facebook, Instagram, Twitter, Snapchat, TikTok, and YouTube. These platforms are categorized based on their primary functions and the types of content they support. Social networks like Facebook, LinkedIn, NextDoor, and X (formerly Twitter) focus on connecting people and building relationships. Media-sharing networks such as Instagram, Pinterest, Snapchat, TikTok, and YouTube allow users to upload diverse media types like photos, videos, and live streams. Messaging networks such as WhatsApp, Signal, Telegram, and Messenger facilitate communication through text, video, or voice. Meanwhile, discussion networks like Reddit, Quora, Stack Exchange, and Discord enable users to engage in online discussions and debates. With the expansion of these platforms, a complex network of user interactions and content choices has evolved. In addition, the availability of a variety of social networking sites, each with its own specific features, has led users to use multiple platforms. Understanding user behavior and preferences is crucial in today's ever-changing social media landscape, influencing scholarly research and practical applications in commerce, marketing (Li et al., 2023; Goyal et al., 2021), and user experience design (Shahbaznezhad et al., 2021). However, to the best of our knowledge, prior studies have not investigated the multiplatform use of social media.

This document introduces a novel method to decipher these complex patterns of social media use by applying Market Basket Analysis (MBA), a strategy typically used in retail sales, to study social media usage. Market Basket Analysis is a data mining technique that discovers relationships and associations between various items. In retail, it helps identify products that are frequently bought together, aiding in marketing strategies, inventory management, and predicting customer purchases (Pang et al., 2020; Said et al., 2012). MBA is a well-established analytical technique commonly applied in the retail sector to inform marketing strategies and product design. However, this method has not been similarly adopted for developing marketing strategies for social media or in the architectural design of social media platforms. Therefore, this paper explains how MBA concepts can be adapted to analyze social media data, treating user actions, engagements on platforms, and content choices as items in an online shopping basket.

The rationale for using the MBA approach to analyze social media data is based on the similarities between shopping behavior in retail and user activities on social media platforms. Just as a shopper's selections in a basket can reveal their preferences and trends, a social media user's platform choice can provide significant insights into their social media preferences and behaviors. This study aims to uncover hidden trends in social media usage, offering valuable insights for platform developers, marketers, and researchers. These insights can help create more engaging and personalized user experiences and offer a fresh perspective for understanding social media dynamics, contributing to the fields of digital sociology and media studies. The following sections will discuss how applying Market Basket Analysis to social media data can provide key insights and examine the implications of these findings in the context of current social media trends and future research directions.

Background Literature

Prior Research

Knowledge discovery involves extracting patterns that represent non-trivial, implicit, previously unknown, but easily understandable and actionable knowledge (Olson, 2017). Association rule mining, a longstanding knowledge discovery technique, has been extensively studied (Aggarwal et al., 1993) and is utilized to extract relevant information from large transactional databases. For example, a transactional database could be a shopping basket database, where the items are products, or a text database where entities are represented by words. Market Basket Analysis (MBA), also known as association rule mining or affinity analysis, originally developed in marketing, has recently been applied effectively across various fields.

In bioinformatics, MBA has uncovered patterns in genes, proteins, and disease associations, and identified drug-effective molecular structures (Naulaerts et al., 2015). It has also predicted nuclear event likelihoods and optimal reactor configurations (Hibino & Niwa, 2008). Aguinis et al. (2013) highlight MBA's ability to support inductive theorizing without conventional statistical assumptions, utilizing even "unusable" data to bridge the gap between science and practice. Additionally, a recent study applied MBA to Google Trends on women's entrepreneurship, identifying public opinion trends and analyzing them through lenses of gender equality and social justice, demonstrating MBA's broad applicability (Semerci et al., 2022). While MBA has been utilized as a tool for knowledge discovery in various fields, its application in uncovering patterns of social media usage remains unexplored. This paper aims to broaden the use of MBA within the realm of social media studies.

Even though MBA hasn't been used to analyze social media data, other association rule mining methods have found significant applications in social media research. Unsupervised data mining methods have been employed to discover associations and extract knowledge from user-generated, unstructured textual entities like tweets (Diaz-Garcia et al., 2023). Association rule mining tasks in social media include summarization

(Phan et al., 2018), event and topic detection (Margono et al., 2014; Zainol et al., 2018), sentiment analysis (Diaz-Garcia et al., 2019), forecasting (Shen et al., 2018; Tundis et al., 2019), and developing recommender systems (Kakulapati & Reddy, 2019; Rao et al., 2019). Previous research has not explored social media data to analyze usage patterns across multiple platforms. Our study employs Market Basket Analysis (MBA) as the method for association rule mining, analyzing survey data on social media use and identifying sets of platforms that are often used in conjunction. The methodology section of this paper details some of the commonly used MBA metrics and how they could be adapted for social media data to identify multiplatform use.

Market Basket Analysis

Market basket analysis (MBA) is based on association rules, which is commonly used to track how often items are purchased together in transactions. For example, if a customer buys bread, they are more likely to also buy butter. This relationship can be expressed as A⇒B, where "A" (the antecedent) is an item like bread found in the basket, and "B" (the consequent) is an item like butter that often accompanies it. These rules might include multiple antecedents and consequents. Association rules are mined from transaction data, where each transaction, or "basket," is identified by a unique ID. The rules are then evaluated using metrics such as support, confidence, lift, and deployability to determine their strength and relevance. Table 1 summarizes these metrics.

Support measures the frequency of item occurrences in the dataset. It can be split into antecedent support and rule support. Antecedent support is the percentage of transactions containing the antecedent, while rule support is the percentage containing both the antecedent and consequent. For instance, if 50% of the data contains item A, and 20% contains both A and B, the antecedent support for A is 50%, and the rule support for A⇒B is 20%.

Confidence measures the likelihood that a transaction with the antecedent will also have the consequent. It is calculated by dividing the rule support by the antecedent support. Using the previous example, the confidence for A⇒B would be 40%, indicating the strength of the implication.

Lift compares the confidence of the rule to the baseline probability of finding the consequent in the dataset. It helps identify the most interesting rules. A lift greater than 1 suggests that the items are more likely to be purchased together than separately. While support and confidence are useful for identifying frequent item sets and strong rules, lift is crucial for understanding whether the association between items is significant or just a random occurrence.

Deployability reflects the percentage of transactions that contain the antecedent but not the consequent. It helps identify potential opportunities for promoting the consequent to customers who have bought the antecedent. Deployability shows the potential for a rule support that is currently not met. These metrics are often used in tandem to ensure that the rules generated from market basket analysis are both frequent and meaningful.

Table 1: Support, Confidence and Lift for the rule A => B

Instances = number of unique ids for which given antecedent(s) is (are) true
(*Antecedent*) Support(A)= percentage of unique IDs that have an antecedent A, or (*frequency of A*)
Rule Support (A, B) = percentage of unique IDs that have both A and B, or (*frequency of A,B*)
Confidence (A, B) = Rule Support/Antecedent Support , or (*frequency* (A, B)/ *frequency*(A))
Lift (A, B) = Rule Support(A, B)/ (Support(A)*Support(B)) = Confidence(A,B)/Support (B)
Deployability (A,B) = (# of Antecedent Support (A) - # of Rule Support (A,B)) / Total number of uniqueIDs or transactions.

Methodology

Data

Data for this study is obtained from a Pew Research report that is based on a survey conducted through phone interviews with 1,502 adults aged 18 or older, living across all 50 U.S. states and the District of Columbia. Abt Associates managed the interview process, during which 300 respondents were contacted via landline and 1,202 via cellphone, including 845 who did not have landlines. The sample, which included both landline and cellphone numbers, was provided by Dynata according to specifications from Abt Associates. The interviews were carried out in both English and Spanish (Pew Research Survey, 2021). For more detailed information on the survey methodology, refer to the U.S. Survey published by the Pew Research Center (Pew Research Survey, 2021).

For the landline portion, the youngest adult male or female present was chosen randomly. For the cellphone portion, the survey was conducted with any adult who answered the phone. The data from both landline and cellphone respondents were then weighted. This weighting process used an iterative method to align the sample's demographic characteristics—gender, age, education, race, Hispanic origin, nativity, region, and population density—with the U.S. Census Bureau's 2019 American Community Survey and the decennial census. Additionally, adjustments were made to reflect current telephone usage patterns based on findings from the 2019 National Health Interview Survey.

Association Rule Discovery

In the context of association rule mining, the variables are the items or attributes in the dataset scrutinized to uncover co-occurrence patterns. Our study focuses on exploring usage patterns across various social media platforms. In this regard, our research identifies eleven social media platforms as the variables, or 'items,' based on their reported usage by individuals in a survey. We utilize these variables to establish rules that illustrate the frequency with which different social media platforms are used in conjunction. The analysis of our Pew dataset MBA was conducted using SPSS Modeler 18.4. Throughout our analysis, we consider each social media platform as an individual variable or 'item,' and the derived association rules reveal the combinations in which these platforms are commonly used.

In this study, we employ the Apriori algorithm to discover the association rules for the social media platforms. The strength of the association rules is evaluated using metrics such as support, confidence, lift, and deployability. These metrics help evaluate the quality and usefulness of the discovered rules, which is crucial for effective data analysis. The Apriori algorithm (Aggarwal & Srikanth, 1994), a well-known method for mining associations, operates under the principle that if an item set is frequent, then all its subsets are also frequent. This allows us to prune the search tree by skipping infrequent item sets. Alternative algorithms designed to handle larger computational burdens include FP-Growth (Han et al., 2000) and Eclat (Zaki et al., 1997). Eclat adopts a depth-first search, focusing on intersections between item lists, which can be memory-intensive. In contrast, FP-Growth uses a compressed database representation through an FP-tree and a selective strategy that increases efficiency when handling large data volumes. For our purposes, Apriori, which is an easy-to-understand algorithm, could easily handle the computational burden caused by our dataset.

Results

The MBA was performed using SPSS Modeler 18.4. The model is shown in Figure 1. The Pew Social Media survey was used as the data source, and variables were selected and typed in the second stage. The

11 fields selected were then processed using the Apriori algorithm and the summary results are shown in Table 2. The named icons shown in Figure 1 depicts the steps in the Apriori model.

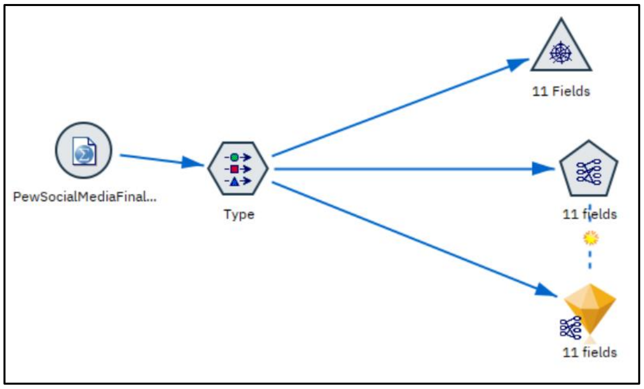

Figure 1: Pew Social Media Apriori model

The model uses fields that could be an input, which are predictors for the model, or they could be outputs, which are targets that the model will try to predict. The model could also use fields as both inputs and outputs, as is the case in this study. For our model, the fields consisted of all social media questions on usage of the 11 platforms from the Pew research data and these fields were coded as both inputs and outputs. The results of the Apriori model are illustrated by the last three icons in Figure 1. The triangle depicts the web graph of relationships. The pentagon indicates the model parameter settings. We primarily used default parameter settings. Finally, the diamond icon shows the display model results depicted as various tables in this section.

Table 2: Summary of Analysis

Number of Rules:	7,007	**Minimum Confidence:**	17.37%	**Minimum Deployability:**	0.00%
Number of Valid Transactions:	1,502	**Maximum Confidence:**	100.00%	**Maximum Deployability:**	66.18%
Minimum Support:	0.399%	**Minimum Lift:**	1.028%	**Minimum Rule Support:**	0.399%
Maximum Support:	80.09%	**Maximum Lift:**	5.51%	**Maximum Rule Support:**	58.32%

Table 2 summarizes the results of our analysis. Please refer to Table 3 for more details of our analysis. Maximum value of support was 81% with YouTube, that can be seen in Table 3, as well. This equates to 81% of respondents using YouTube. With our maximum number of antecedents set at the default number of 5, minimum support and rule support with YouTube was for the combination of Nextdoor, Reddit, TikTok, What's App and Twitter and the Consequent YouTube. The maximum rule support was for YouTube and Facebook with 58.322%. This says that YouTube and Facebook are the most popular combination. Maximum lift value that could not be shown in Appendix A due to page limitation, was the combination of Pinterest, Snapchat, Reddit, Twitter and Nextdoor as antecedents for TikTok usage. Users of Pinterest, Snapchat, Reddit, Twitter and Nextdoor are more than 5 times more likely to be TikTok users than random. Confidence levels maximums of 100% were reached with a multitude of platform configurations as antecedents and YouTube, Facebook, and Instagram as consequents. The highest deployability were between YouTube and Nextdoor, TikTok, Reddit, Snapchat, WhatsApp, and Twitter. As noted, deployability reflects the percentage of transactions that contain the antecedent but not the consequent. It helps identify potential opportunities for promoting the consequent to customers who have bought the antecedent.

Table 3 : Association Rules and their metrics, sorted by the value of lift

# Links	Consequent	Antecedent	Instances	Antecedent Support %	Confidence %	Rule Support %	Lift	Deployability
244	Instagram	Snapchat	307	20.4	79.5	16.2	2.3	4.2
154	Twitter	Snapchat	307	20.4	50.2	10.3	2.2	10.2
187	Instagram	TikTok	248	16.5	75.4	12.5	2.1	4.1
254	Twitter	Instagram	530	35.3	47.9	16.9	2.1	18.4
190	WhatsApp	LinkedIn	480	32.0	39.6	12.6	1.8	19.3
202	Twitter	LinkedIn	480	32.0	42.1	13.4	1.8	18.5
151	LinkedIn	Reddit	260	17.3	58.1	10.1	1.8	7.3
161	Instagram	Reddit	260	17.3	61.9	10.7	1.8	6.6
185	Instagram	WhatsApp	325	21.6	56.9	12.3	1.6	9.3
235	Instagram	Pinterest	436	29.0	53.9	15.6	1.5	13.4
257	Instagram	LinkedIn	480	32.0	53.5	17.1	1.5	14.8
192	Pinterest	LinkedIn	480	32.0	40.0	12.8	1.4	19.2
373	Facebook	Pinterest	436	29.0	85.6	24.8	1.3	4.2
449	Instagram	Facebook	988	65.8	45.4	29.9	1.3	35.9
292	Twitter	Facebook	988	65.8	29.6	19.4	1.3	46.3
208	Facebook	TikTok	248	16.5	83.9	13.8	1.3	2.7
256	Facebook	Snapchat	307	20.4	83.4	17.0	1.3	3.4
267	Facebook	WhatsApp	325	21.6	82.2	17.8	1.2	3.9
177	Facebook	Nextdoor	220	14.6	80.5	11.8	1.2	2.9
337	Twitter	YouTube	1203	80.1	28.0	22.4	1.2	57.7
251	YouTube	Reddit	260	17.3	96.5	16.7	1.2	0.6
239	YouTube	TikTok	248	16.5	96.4	15.9	1.2	0.6
379	Facebook	LinkedIn	480	32.0	79.0	25.2	1.2	6.7
461	YouTube	LinkedIn	480	32.0	96.0	30.7	1.2	1.3
506	Instagram	YouTube	1203	80.1	42.1	33.7	1.2	46.4
209	YouTube	Nextdoor	220	14.6	95.0	13.9	1.2	0.7
291	Snapchat	YouTube	1203	80.1	24.2	19.4	1.2	60.7
306	YouTube	WhatsApp	325	21.6	94.2	20.4	1.2	1.3
406	YouTube	Pinterest	436	29.0	93.1	27.0	1.2	2.0
191	Facebook	Reddit	260	17.3	73.5	12.7	1.1	4.6
876	Facebook	YouTube	1203	80.1	72.8	58.3	1.1	21.8

Table 3 shows the distribution of the number of social media users who have reported that they use two social media platforms. Unlike support, which merely highlights popular items or item combinations, lift considers both the support of individual items and their joint occurrence, providing a more comprehensive understanding of relationships. Furthermore, lift surpasses confidence by accounting for the expected co-occurrence of platforms based on their individual support levels. This enables businesses to prioritize their efforts on item combinations that exhibit a significant lift value, steering them away from wasting resources on items with high confidence but weak associations.

Table 3 shows higher lift values of greater than 2.0 for the rules Snapchat=>Instagram, Snapchat=>Twitter, Tik-Tok => Instagram and Instagram=>Twitter. These rules do not necessarily indicate highest confidence or rule support. The rule YouTube=>Snapchat displays highest deployability, followed by YouTube => Twitter. When additional antecedents were added to the Instagram and YouTube, the lift values improved. Table 4 shows how adding additional social media platforms as antecedents improves the lift of the association rules with Instagram and Facebook consequents.

Table 4: An example showing association rule strength for Instagram and Facebook as consequents.

Consequent	Antecedent	Rule Support %	Confidence %	Lift
Instagram	YouTube	42.1	33.7	1.2
Instagram	Snapchat	79.5	16.2	2.3
Instagram	Facebook	45.4	29.9	1.3
Instagram	YouTube, Snapchat and Twitter	10.12	93.42	2.65
Instagram	YouTube, Snapchat and Facebook	16.11	82.23	2.33
Instagram	YouTube and Snapchat	19.37	80.07	2.27
Facebook	YouTube	72.8	58.3	1.1
Facebook	Pinterest	85.6	24.8	1.3
Facebook	YouTube, Pinterest and Instagram	14.91	92.86	1.41
Facebook	YouTube, Pinterest and LinkedIn	12.58	92.06	1.40

The results of running the Apriori algorithm for a given consequent, over possible antecedents are shown in the table in Appendix A. Only those association rules, whose lift is greater than 1.0 are displayed in the table. This table shows the item sets that have Instagram as a consequent, also had highest value of lift compared to several other significant consequents such as Facebook and YouTube.

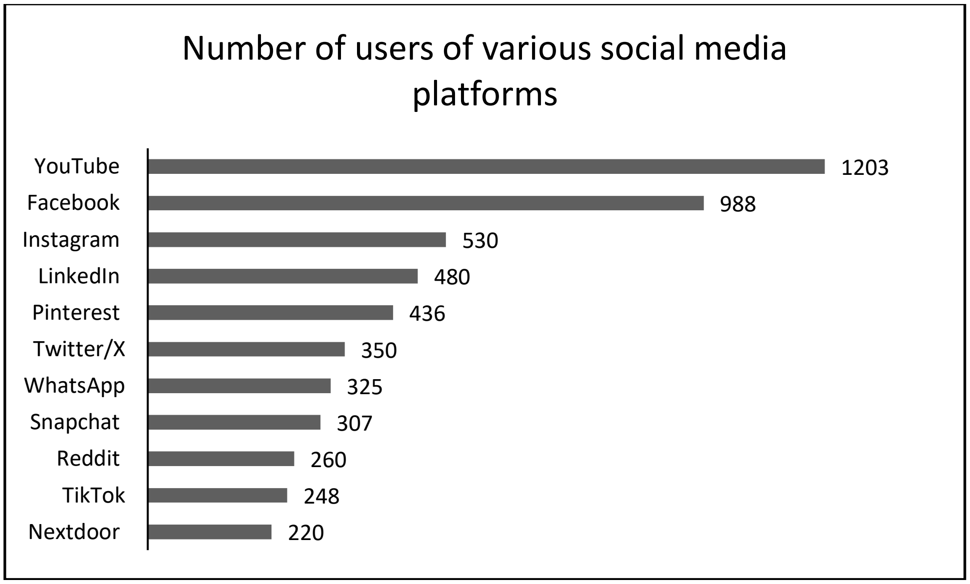

Figure 2: Popularity of social media usage among the survey respondents

Figure 2 shows that Instagram, Facebook and YouTube were also among the platforms that were most used by the respondents. Combinations of any two of these platforms also had many users, as shown by the "# of Links" columns in Table 3. However, higher predictability of Instagram as a consequent required the antecedent Snapchat that nearly did not have many users. Similarly, predictability of Facebook, or YouTube, as a consequent required a combination of social media platforms as antecedents that included other lesser popular platforms.

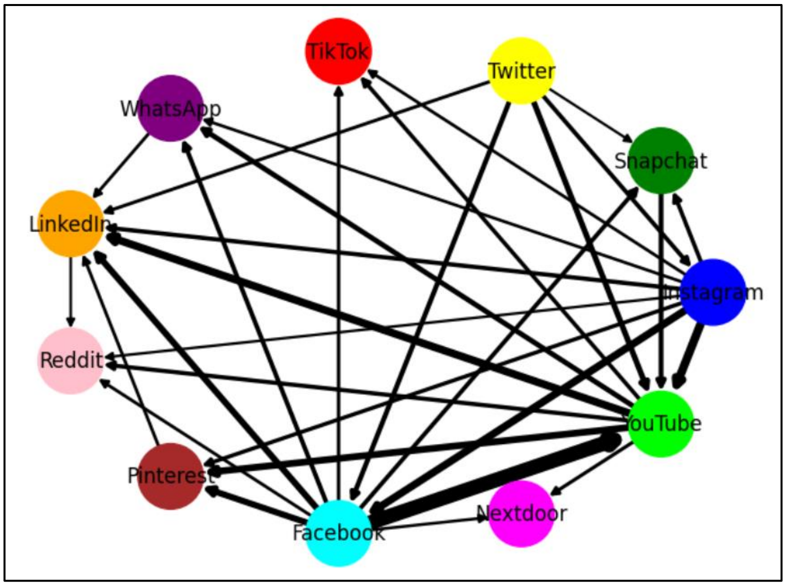

Figure 3: A web chart of the associations among social media use.

Figure 2 shows the distribution of survey respondents who reported using various social media platforms. YouTube, Facebook and Instagram were the top three popular social media platforms. A web chart of associations, as shown in Figure 3, indicates relatively strong links between Facebook, YouTube and Instagram.

Discussion

The ability to predict multi-platform usage through association rules offers significant advantages for platform designers, businesses, and organizations. Our analysis identifies Facebook, Instagram, and YouTube as key consequents in strong association rules, with antecedents of varying popularity. Such robust connections between these social media platforms can facilitate the flow of traffic and content from one to another, providing substantial marketing benefits and opening new avenues for creative content strategies and platform design. Metrics such as lift can be utilized to predict highly probable consequents, while deployability can be employed to identify potential improvements in platform design associations. This section explores some of the potential application scenarios for Market Basket Analysis (MBA).

Business and Marketing Implications

When used for commerce, social media users act as both consumers and promoters of products. Features such as liking, sharing, and reposting on social media are invaluable tools for assessing user engagement, content quality, and perceived value. These features also serve as key performance indicators for content strategies aimed at disseminating information and engaging users across social media platforms (Kumar, 2013).

Marketing strategies on social media largely employ the "word of mouth" (WOM) approach in various forms. Marketing scholars, including Lamberton and Stephen (2016), have extensively characterized social media in this context. Meta-analytic studies, such as those by Rosario et al., (2016), consistently find a positive correlation between online WOM and sales. Engaging with multiple platforms allows users to share information from one platform to another, enhancing reach and impact. Platforms associations with high lift values could be used for WOM approach to improve sales of products. For example, WOM might be predicted with a very high confidence for the association Snapchat => Instagram (with a lift value of 2.3, as shown in Table 3) than it is for the association Youtube => Facebook. This result cannot be deduced by the results of Figure 2 alone that shows both Youtube and Facebook as highly popular platforms.

Different platforms cater to specific goals. For example, content creators may focus on building awareness on Instagram and Twitter while using Facebook ads to drive conversions. A cross-platform strategy ensures that content is appropriate and customized for each platform where it's posted. Dove's "Real Beauty" campaign is a prime example of effective cross-platform engagement, reaching audiences on Instagram, Pinterest, and YouTube and sparking discussions on self-acceptance (Kramer et al., 2019). Similarly, Coca-Cola's personalized bottle campaign encouraged users to share their unique Coke bottles on social media, generating significant buzz on Instagram, Facebook, and Twitter (Nguyen, 2021). While it is unclear if these campaigns utilized association rule mining to select social media platforms, such strategies underscore the potential of using market basket analysis to uncover strong, predictable rules for multi-platform social media use. Cross-platform campaigns could use lift values from a MBA to predict associations that are highly likely and the target their message across these platforms. For example, Table 4 shows very high lift values for the rule (YouTube, Snapchat, Twitter => Instagram). This means, a cross-platform campaign that target these antecedents are very highly likely to see the message being posted on the consequent, Instagram.

Recognizing strong associations across different platforms enables content developers to create targeted cross-promotion strategies for brands, products, and campaigns. Cross-platform content promotion strategically spreads content across multiple platforms, enhancing engagement. Studies by Unnava and Aravindakshan (2021) demonstrate how consumer engagement evolves with brand posts across multiple platforms. Their findings indicate that brand communications on one platform can lead to 'spillover effects' on others, prolonging engagement, and interaction. These effects, which vary among platforms due to their unique characteristics, can amplify the "virality" of content. Content introduced on a leading platform can spill over to a secondary platform and may even circle back to the initial platform, increasing its impact (Krijestorac et al., 2019).

Effective use of association rule mining can help businesses develop resource allocation models that could predict these spillover and carryover effects, optimizing content virality based on platform characteristics. The consequent platform of strong association rules could be leveraged as candidates of spillover effects with a high predictability. For example, Snapchat => Instagram association with a lift value of 2.3, means spillover effects could be potentially predicted with high confidence for this association. As we may observe from Table 4, targeting more antecedents could improve the lift of the association rules and hence, spillover into the consequent is more likely. For example, picking the association (Snapchat, Youtube, Twitter => Instagram) with a lift of 2.65 and spreading the contents to the three antecedents could predict a higher likelihood of spillover into Instagram.

Implication for Platform Designers

By strategically leveraging the associations between multi-platform social media use, developers can craft cohesive, engaging, and user-centered platforms that stand out in a crowded digital ecosystem.

Understanding the user interface (UI) preferences and navigation habits of multi-platform users allows developers to create more intuitive and user-friendly interfaces. Incorporating familiar elements from other platforms can make users more comfortable, thereby reducing the learning curve and enhancing engagement and retention. For instance, the seamless posting of a TikTok video as an Instagram reel, as demonstrated by the high lift value of the TikTok =>Instagram association rule shown in Table 1, exemplifies this ease of cross-platform sharing.

Analyzing multi-platform association rules provides insights into the competitive landscape, enabling developers to identify successful features and strategies used by competing platforms. This knowledge can drive innovation, prompting developers to devise unique features that address gaps in the user experience. Developers can utilize the deployability metric, which evaluates the potential for rule support between antecedents and consequents, to examine how existing platform designs could be improved to better facilitate cross-platform content sharing. For example, high deployability of Youtube => Snapchat could be a potential design opportunity to include features in Snapchat that may be compatible with Youtube. Additionally, platform developers could design features that foster and support strong associations, benefiting content providers by making the sharing process more efficient and effective.

Conclusion

The application of Market Basket Analysis (MBA) to the exploration of social media data elucidates significant insights into user interaction patterns and platform utilization. By conceptualizing user activities and content preferences as elements within a "shopping basket," this methodology not only reveals platform synergies but also discerns latent behavioral trends among users. Such insights are indispensable for platform developers and marketers aiming to forge more engaging and tailored user experiences. Additionally, the adaptability of MBA across diverse data landscapes—from traditional retail to digital interactions—highlights its robustness and versatility as an analytical tool in the data-intensive environment of today.

The results of this study underscore the critical role of understanding dynamics across multiple social media platforms. They afford a strategic advantage to businesses and content strategists, empowering them to refine their engagement and marketing tactics based on predictive associations between platforms. As the landscape of social media continues to evolve, harnessing the power of such advanced analytics will be paramount in maintaining competitive advantage, enabling stakeholders to deploy targeted and impactful strategies that resonate with an increasingly diverse user base. Continued application and enhancement of data mining techniques like MBA will be essential in anticipating and adapting to the dynamic preferences of digital consumers, thereby aligning strategic objectives with user-centric data insights.

References

Aggarwal, R., Imielinski, T., & Swami, A. (1993). Database mining: A performance perspective. *IEEE Transactions on Knowledge and Data Engineering*, *5*(6), 914-925.

Aggarwal, R., & Srikant, R. (1994). Fast algorithms for mining association rules. In *Proceedings of the 20th International Conference on Very Large Data Bases* (VLDB) (Vol. 1215, pp. 487–499).

Aguinis, H., Forcum, L. E., & Joo, H. (2013). Using Market Basket Analysis in Management Research. Journal of Management, 39(7), 1799-1824. https://doi.org/10.1177/0149206312466147

Rosario, B. A., Sotgiu, F., De Valck, K., & Bijmolt, T. H. A. (2016). The effect of electronic word of mouth on sales: A meta-analytic review of platform, product, and metric factors. Journal of Marketing Research, 53(3), 297–318.

Diaz-Garcia, J. A., Ruiz, M. D., & Martin-Bautista, M. J. (2019). Generalized association rules for sentiment analysis on Twitter. In *International Conference on Flexible Query Answering Systems* (pp. 166–175). Springer.

Diaz-Garcia, J. A., Ruiz, M. D., & Martin-Bautista, M. J. (2023). A survey on the use of association rules mining techniques in textual social media. *Artificial Intelligence Review, 56*, 1175–1200. https://doi.org/10.1007/s10462-022-10196-3

Goyal, S., Hu, C., Chauhan, S., Gupta, P.J., Bhardwaj, A.K., & Mahindroo, A. (2021). Social Commerce: A Bibliometric Analysis and Future Research Directions. *J. Glob. Inf. Manag., 29,* 1-33.

Han, J., Pei, J., & Yin, Y. (2000). Mining frequent patterns without candidate generation. *ACM SIGMOD Record, 29*(1), 1–12.

Hibino, A., & Niwa, Y. (2008). Graphical representation of nuclear incidents/accidents by associating network in nuclear technical communication. Journal of Nuclear Science and Technology, 45(5), 369-377.

Kakulapati, V., & Reddy, S. M. (2019). Mining social networks: Tollywood reviews for analyzing UPC by using big data framework. In *Smart Innovations in Communication and Computational Sciences* (pp. 323–334). Springer.

Kramer, M. R., Sidibe, M., & Veda, G. (2019). Dove and real beauty: Building a brand with purpose. *Harvard Business Review*. Retrieved from https://hbr.org

Krijestorac, H., Garg, R., & Mahajan, V. (2019). Cross-platform spillover effects in consumption of viral content: A quasi-experimental analysis using synthetic controls. Retrieved from https://ssrn.com/abstract=3011533 or http://dx.doi.org/10.2139/ssrn.3011533

Kumar, V., Bhaskaran, V., Mirchandani, R., & Shah, M. (2013). Creating a measurable social media marketing strategy: Increasing the value and ROI of intangibles and tangibles for Hokey Pokey. Marketing Science, 32(2), 194–212.

Lamberton, C., & Stephen, A. T. (2016). A thematic exploration of digital, social media, and mobile marketing research's evolution from 2000 to 2015 and an agenda for future research. Journal of Marketing, 80(6), 146–172.

Li, F., Larimo, J., & Leonidou, L. C. (2023). Social media in marketing research: Theoretical bases, methodological aspects, and thematic focus. Psychology & Marketing, 40, 124–145. https://doi.org/10.1002/mar.21746

Margono, H., Yi, X., & Raikundalia, G. K. (2014). Mining Indonesian cyber bullying patterns in social networks. In *Proceedings of the Thirty-Seventh Australasian Computer Science Conference* (Vol. 147, pp. 115–124). Australian Computer Society, Inc.

Naulaerts, S., Meysman, P., Bittremieux, W., Vu, T. N., Vanden Berghe, W., Goethals, B., & Laukens, K. (2015). A primer to frequent itemset mining for bioinformatics. Briefings in Bioinformatics, 16(2), 216–231. https://doi.org/10.1093/bib/bbt074

Nguyen, N. (2021). Coca Cola's Share a Coke campaign. Retrieved from https://www.smithbrothersmedia.com.au/case-study/coca-colas-share-a-coke-campaign/

Zaki, M. J., Parthasarathy, S., Ogihara, M., & Li, W. (1997). New algorithms for fast discovery of association rules. In *KDD'97: Proceedings of the Third International Conference on Knowledge Discovery and Data Mining* (pp. 283-286).

Olson, D. L. (2017). Market basket analysis. In D. L. Olson (Ed.), *Descriptive Data Mining: Computational Risk Management* (pp. 1-3). Springer, Singapore. https://doi.org/10.1007/978-981-10-3340-7_3

Pang, C., Wu, X., Ji, X., & Quan, V. (2020). Bibliometrics of Social Media Marketing: The Development Status and Quantitative Analysis. *2020 IEEE 11th International Conference on Software Engineering and Service Science (ICSESS)*, 449-452.

Pew Research Center. (2021). Social media use in 2021. Retrieved from https://www.pewresearch.org/internet/2021/social-media-use

Phan, H. T., Nguyen, N. T., & Hwang, D. (2018). A tweet summarization method based on maximal association rules. In *International Conference on Computational Collective Intelligence* (pp. 373–382). Springer.

Rao, P. G., Khan, M. Z., Rafeeq, M., Hitesh, N., Shenoy, P. D., & Venugopal, K. (2019). An automated learning system for Twitter trends. In *2019 Fifteenth International Conference on Information Processing* (ICINPRO) (pp. 1–6). IEEE.

Said, A., Panneer Selvam, D. D. D., & Zailani, S. (2012). A new scheme for extracting association rules: Market basket analysis case study. *International Journal of Business Innovation and Research, 6*, 28-46. https://doi.org/10.1504/IJBIR.2012.044256

Semerci, A. B., Özgören, A. A., & İçen, D. (2022). Thoughts on women entrepreneurship: An application of market basket analysis with Google Trends data. *Soft Computing*, *26*, 10035–10047. https://doi.org/10.1007/s00500-022-07355-7

Shahbaznezhad, H., Dolan, R., & Rashidirad, M. (2021). The role of social media content format and platform in users' engagement behavior. Journal of Interactive Marketing, 53(1), 47-65. https://doi.org/10.1016/j.intmar.2020.05.001

Shen, D., Zhang, L., Cao, J., & Wang, S. (2018). Forecasting citywide traffic congestion based on social media. *Wireless Personal Communications, 103*(1), 1037–1057.

Tundis, A., Jain, A., Bhatia, G., & Muhlhauser, M. (2019). Similarity analysis of criminals on social networks: An example on Twitter. In *2019 28th International Conference on Computer Communication and Networks* (ICCCN) (pp. 1–9). IEEE.

Unnava, V., & Aravindakshan, A. (2021). How does consumer engagement evolve when brands post across multiple social media? Journal of the Academy of Marketing Science, 49, 864–881. https://doi.org/10.1007/s11747-021-00785-z

Zainol, Z., Wani, S., Nohuddin, P. N., Noormanshah, W. M., & Marzukhi, S. (2018). Association analysis of cyberbullying on social media using Apriori algorithm. *International Journal of Engineering & Technology, 7*(4.29), 72–75.

Appendix A

Table A – Market Basket Analysis Results showing associations that have a lift greater than 1.0

Consequent	Antecedent	Support %	Confidence %	Lift
Instagram	Snapchat and Twitter	10.25	93.51	2.65
Instagram	Snapchat and Twitter and YouTube	10.12	93.42	2.65
Instagram	Snapchat and Facebook and YouTube	16.11	82.23	2.33
Instagram	Snapchat and Facebook	17.04	81.25	2.30
Instagram	Snapchat and YouTube	19.37	80.07	2.27
Facebook	Pinterest and Instagram and YouTube	14.91	92.86	1.41
Facebook	Pinterest and Instagram	15.65	92.77	1.41
Facebook	Pinterest and LinkedIn	12.78	92.19	1.40
Facebook	Pinterest and LinkedIn and YouTube	12.58	92.06	1.40
Facebook	WhatsApp and Instagram and YouTube	11.98	90.00	1.37
Facebook	WhatsApp and Instagram	12.32	89.73	1.36
Facebook	Twitter and LinkedIn and Instagram	10.19	89.54	1.36
Facebook	Twitter and LinkedIn and Instagram and YouTube	10.05	89.40	1.36
Facebook	Twitter and Instagram and YouTube	16.71	88.84	1.35
Facebook	Twitter and Instagram	16.91	88.58	1.35
Facebook	Snapchat and Twitter and YouTube	10.12	88.16	1.34
Facebook	Snapchat and Twitter	10.25	87.66	1.33
Facebook	LinkedIn and Instagram	17.11	87.55	1.33
Facebook	LinkedIn and Instagram and YouTube	16.84	87.35	1.33
Facebook	Twitter and LinkedIn and YouTube	13.18	86.87	1.32
Facebook	TikTok and Instagram and YouTube	12.12	86.81	1.32
Facebook	Twitter and LinkedIn	13.45	86.63	1.32
Facebook	TikTok and Instagram	12.45	86.63	1.32
Facebook	WhatsApp and LinkedIn and YouTube	12.38	86.02	1.31
Facebook	Pinterest and YouTube	27.03	85.96	1.31
Facebook	WhatsApp and LinkedIn	12.65	85.79	1.30
Facebook	Pinterest	29.03	85.55	1.30
Facebook	Snapchat and Instagram and YouTube	15.51	85.41	1.30
Facebook	Snapchat and Instagram	16.25	85.25	1.30
Facebook	Instagram and YouTube	33.69	84.98	1.29
Facebook	Twitter and YouTube	22.44	84.87	1.29
Facebook	Instagram	35.29	84.72	1.29
Facebook	Twitter	23.04	84.39	1.28
Facebook	TikTok and YouTube	15.91	84.10	1.28
Facebook	WhatsApp and YouTube	20.37	83.99	1.28
Facebook	TikTok	16.51	83.87	1.28
Facebook	Reddit and Instagram and YouTube	10.32	83.87	1.28
Facebook	Reddit and Instagram	10.72	83.85	1.27

Consequent	Antecedent	Support %	Confidence %	Lift
Facebook	Snapchat	20.44	83.39	1.27
Facebook	Snapchat and YouTube	19.37	83.16	1.26
Facebook	WhatsApp	21.64	82.15	1.25
YouTube	Twitter and Instagram and Facebook	14.98	99.11	1.24
YouTube	Twitter and Instagram	16.91	98.82	1.23
YouTube	Snapchat and Twitter	10.25	98.70	1.23
YouTube	Twitter and LinkedIn and Instagram	10.19	98.69	1.23
YouTube	Reddit and LinkedIn	10.05	98.68	1.23
Facebook	Nextdoor and YouTube	13.91	80.86	1.23
YouTube	LinkedIn and Instagram	17.11	98.44	1.23
YouTube	Pinterest and LinkedIn	12.78	98.44	1.23
YouTube	Pinterest and LinkedIn and Facebook	11.78	98.31	1.23
YouTube	Twitter and LinkedIn and Facebook	11.65	98.29	1.23
YouTube	LinkedIn and Instagram and Facebook	14.98	98.22	1.23
YouTube	WhatsApp and LinkedIn and Facebook	10.85	98.16	1.23
YouTube	Twitter and LinkedIn	13.45	98.02	1.22
Facebook	Nextdoor	14.65	80.45	1.22
YouTube	Twitter and Facebook	19.44	97.95	1.22
YouTube	WhatsApp and LinkedIn	12.65	97.89	1.22
YouTube	WhatsApp and Instagram and Facebook	11.05	97.59	1.22
YouTube	TikTok and Instagram and Facebook	10.79	97.53	1.22
YouTube	Twitter	23.04	97.40	1.22
YouTube	TikTok and Instagram	12.45	97.33	1.22
YouTube	WhatsApp and Instagram	12.32	97.30	1.21
YouTube	Reddit and Facebook	12.72	96.86	1.21
YouTube	LinkedIn and Facebook	25.23	96.83	1.21
YouTube	TikTok and Facebook	13.85	96.63	1.21
YouTube	Reddit	17.31	96.54	1.21
YouTube	TikTok	16.51	96.37	1.20
YouTube	Reddit and Instagram	10.72	96.27	1.20
YouTube	WhatsApp and Facebook	17.78	96.25	1.20
YouTube	LinkedIn	31.96	96.04	1.20
YouTube	Instagram and Facebook	29.89	95.77	1.20
YouTube	Snapchat and Instagram and Facebook	13.85	95.67	1.19
YouTube	Snapchat and Instagram	16.25	95.49	1.19
YouTube	Nextdoor and Facebook	11.78	95.48	1.19
YouTube	Instagram	35.29	95.47	1.19
YouTube	Pinterest and Instagram and Facebook	14.51	95.41	1.19
YouTube	Pinterest and Instagram	15.65	95.32	1.19
YouTube	Nextdoor	14.65	95.00	1.19
YouTube	Snapchat	20.44	94.79	1.18

Consequent	Antecedent	Support %	Confidence %	Lift
YouTube	Snapchat and Facebook	17.04	94.53	1.18
YouTube	WhatsApp	21.64	94.15	1.18
YouTube	Pinterest and Facebook	24.83	93.57	1.17
YouTube	Pinterest	29.03	93.12	1.16
YouTube	Facebook	65.78	88.66	1.11