D212: Association Rules and Lift Analysis

Telecommunication Churn Data

Darien Nguyen

Western Governors University

D212 Task Three: Clustering Techniques

June 21, 2023

**Part I: Research Question**

A.  Describe the purpose of this data mining report by doing the following:

1.  Propose **one** question relevant to a real-world organizational situation that you will answer using market basket analysis.   
  
Using market basket analysis, what is the correlation of services/products associated with customer churn?

2.  Define **one** goal of the data analysis. Ensure that your goal is reasonable within the scope of the scenario and is represented in the available data.

Identifying the services/products that are strongly associated with customer churn and develop targeted strategies to reduce churn rates.

**Part II: Market Basket Justification**

B.  Explain the reasons for using market basket analysis by doing the following:

1.  Explain how market basket analyzes the selected dataset. Include expected outcomes.

Market basket analysis utilizes association rule mining to uncover relationships and

patterns within a dataset. This technique helps identify correlations and dependencies between different items in transactional or relational databases. By analyzing the presence and co-occurrence of items, association rules are generated to capture these associations.

One of the key benefits of association rule mining is gaining insights into customer preferences and behavior. By exploring patterns in service subscriptions, you can understand the characteristics and preferences of different customer segments. For example, you may discover that younger customers with higher incomes are more inclined to choose fiber optic internet service. This knowledge empowers you to make informed decisions and tailor your offerings to specific customer segments.

Moreover, association rule mining enhances marketing and business strategies. By leveraging the discovered associations, you can customize your marketing campaigns, promotions, and product offerings to target specific customer segments. For instance, if customers with device protection services tend to have higher monthly charges, you can create targeted offers for similar customers. This targeted approach improves the effectiveness of your strategies and increases customer engagement.

Furthermore, understanding the relationships between services contributes to improved customer satisfaction and retention. By identifying patterns in service subscriptions, you can provide personalized recommendations to customers. This enhances their overall experience and increases satisfaction. Additionally, you can optimize customer support strategies and develop retention initiatives based on the associations uncovered through the analysis. This proactive approach helps to foster long-term relationships with customers.

2.  Provide **one** example of transactions in the dataset.

3.  Summarize **one** assumption of market basket analysis.

Closed-world assumption: Market basket analysis assumes a closed-world scenario, where all possible item combinations are known and accounted for. This means that associations are derived from the observed itemsets within the dataset, and any unobserved itemsets are considered irrelevant. In practice, there may be undiscovered item associations that are not present in the dataset.

**Part III: Data Preparation and Analysis**

C.  Prepare and perform market basket analysis by doing the following:

1.  Transform the dataset to make it suitable for market basket analysis. Include a copy of the cleaned dataset.

2.  Execute the code used to generate association rules with the Apriori algorithm. Provide screenshots that demonstrate the error-free functionality of the code.

A screenshot of a computer program

Description automatically generated with low confidence

3.  Provide values for the support, lift, and confidence of the association rules table.

The association rules table generated by the Apriori algorithm provides insights into the relationships between different items in the dataset. In this case, we have identified the top three rules based on their support, lift, and confidence values.

The first rule, (Churn, Phone), has a support of 1.0, indicating that both "Churn" and "Phone" appear together in all transactions. The confidence is also 1.0, meaning that whenever "Churn" is present, "Phone" is always associated with it. The lift value of 1.0 suggests that the occurrence of "Phone" is independent of "Churn".

Similarly, the second rule, (Phone, Churn), also has a support, confidence, and lift value of 1.0. This implies a strong relationship between "Phone" and "Churn" where the presence of one implies the presence of the other.

The third rule, (Churn, Multiple), follows the same pattern with a support, confidence, and lift value of 1.0. It indicates that whenever "Churn" is observed, "Multiple" is always present as well.

4.  Identify the top **three** rules generated by the Apriori algorithm. Include a screenshot of the top rules along with their summaries.

The Apriori algorithm generated several rules based on the analyzed dataset. Among them, the top three rules are particularly interesting. The first rule suggests a strong association between the "Churn" and "Phone" variables, indicating that customers who churn are likely to have phone-related interactions. The second rule implies a reciprocal relationship, stating that customers with phone-related interactions are prone to churn. The third rule highlights a connection between "Churn" and "Multiple," suggesting that customers who churn may have multiple interactions. These insights provide valuable information for understanding customer behavior and can guide strategic decision-making in customer retention and service improvements.

A screenshot of a computer

Description automatically generated with medium confidence

**Part IV: Data Summary and Implications**

D.  Summarize your data analysis by doing the following:

1.  Summarize the significance of support, lift, and confidence from the results of the analysis.

The significance of support, lift, and confidence in market basket analysis cannot be understated. These metrics provide valuable insights into the relationships between items within a dataset, shedding light on patterns and associations that can drive business decisions.

Support refers to the frequency or occurrence of a specific itemset in the dataset. In our analysis, a support value of 1.0 for all the rules suggests that the antecedent and consequent items appear together in every transaction. This indicates a strong association between these items and implies that they are often purchased together. The high support values highlight the popularity and co-occurrence of these items, which can inform decisions related to product placement, inventory management, and promotional strategies.

Confidence, on the other hand, measures the conditional probability of finding the consequent items given the presence of the antecedent items. In our analysis, the confidence values of 1.0 indicate that whenever the antecedent items are present, the consequent items are always observed as well. This high level of confidence suggests a strong dependency between the items, indicating that the presence of one item significantly influences the purchase of the other. Businesses can leverage this insight to design effective cross-selling campaigns and personalized recommendations to enhance customer satisfaction and increase sales.

Lift is a measure of how much more likely the occurrence of the consequent items is, given the presence of the antecedent items, compared to their individual probabilities. The lift values of 1.0 in our analysis indicate that there is no significant association between the antecedent and consequent items beyond what would be expected by chance. While this may imply independence between the items, it also suggests a lack of meaningful patterns or associations in the dataset. Further analysis and exploration may be necessary to uncover more interesting and actionable insights.

2.  Discuss the practical significance of the findings from the analysis.

The findings from the market basket analysis have practical significance for the business in various ways. Firstly, identifying the strong associations between items can help our telecom company optimize their product placement and store layout. By strategically promoting deals can encourage customers to make additional purchases, increasing sales and enhancing the overall shopping experience.

Secondly, the analysis highlights the potential for cross-selling opportunities. When certain items are frequently purchased together, we can create bundled offers or promotions to encourage customers to buy complementary products. This can lead to increased customer satisfaction and loyalty, as well as higher average transaction values.

Moreover, understanding the relationships between items enables us to personalize their marketing efforts. By leveraging the knowledge of frequently associated items, the telecom company can develop targeted marketing campaigns and recommendations. This personalized approach can enhance the customer experience, increase engagement, and ultimately drive higher conversion rates.

Additionally, the analysis can inform inventory management decisions. By identifying popular item combinations, businesses can ensure sufficient stock levels to meet customer demand. This helps to minimize out-of-stock situations and potential revenue loss. Moreover, businesses can optimize their procurement processes by leveraging insights into the association patterns and adjusting their supply chain accordingly.

Furthermore, the analysis can aid in strategic decision-making, such as product development and assortment planning. By understanding the relationships between items, we can identify gaps in their product offerings or potential opportunities for expanding their product lines. This can lead to diversification, increased market share, and a competitive advantage in the industry.

3.  Recommend a course of action for the real-world organizational situation from part A1 based on your results from part D1.

Based on the results of the market basket analysis, I recommend the following two course of actions for the company:

1. Product Placement Optimization: Utilize the insights from the analysis to strategically place related items in close proximity within the physical store or on the website. By grouping and displaying frequently associated items together, you can encourage customers to make additional purchases. This can lead to increased sales and improved customer satisfaction. For example, if the analysis reveals a strong association between smartphones and phone accessories, you can position phone cases, screen protectors, and chargers near the smartphone display to prompt customers to purchase these complementary items.
2. Cross-Selling and Bundling: Take advantage of the identified associations by creating bundled offers or promotions. Encourage customers to buy related or complementary products together by providing discounts or incentives. This cross-selling strategy can increase the average transaction value and drive higher revenue. For instance, if the analysis reveals a strong association between coffee machines and coffee capsules, you can offer a discounted bundle that includes both items. This can incentivize customers to purchase both products simultaneously and increase overall sales.

By implementing these two recommendations, the company can capitalize on the associations between items to optimize sales and enhance the customer shopping experience. These strategies can be implemented both in physical stores and online platforms. It is essential to monitor and evaluate the effectiveness of these actions to ensure their impact on sales and customer satisfaction and make adjustments as needed based on customer feedback and market trends.

**Part V: Attachments**

E.  Provide a Panopto video recording that includes a demonstration of the functionality of the code used for the analysis and a summary of the programming environment.

<https://wgu.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=375f3d97-ee8f-49bf-973e-b02f006fb512#>

F.  Record *all* web sources used to acquire data or segments of third-party code to support the application. Ensure the web sources are reliable.  
No Sources used.

G.  Acknowledge sources, using in-text citations and references, for content that is quoted, paraphrased, or summarized.  
No sources used.

H.  Demonstrate professional communication in the content and presentation of your submission.