Time Series Analysis of Customer Churn Data

Mike Mattinson

Western Governors University

Advanced Data Analytics – D213

Task 1: Time Series Analysis

Dr. Sewell

May 28, 2022

Revision 1

Abstract

Transaction data will be analyzed using Apriori algorithm also known as “Market Basket Analysis” or “Association Rules”. Approximately 7,500 customer transactions will be considered. There are approx. 119 unique items used in those transactions. The top association rules will be generated and reviewed. Minimum support of 3/1000. Minimum confidence of 3/10. Rules with lift values above 1 are considered to have strong correlation between the antecedents and consequents. Specifically, blue-light reading glasses have a six (6) times likelihood of purchase when computer accessories of “Dust-Off Compressed Gas 2 pack” and “Anker 2-in-1 USB Card Reader” are purchased.

Keywords: Data Mining. Market Basket Analysis. Association Rules and Lift Analysis.

Scenario 1

One of the most critical factors in customer relationship management that directly affects a company’s long-term profitability is understanding its customers. When a company can better understand its customer characteristics, it is better able to target products and marketing campaigns for customers, resulting in better profits for the company in the long term.

You are an analyst for a telecommunications company that wants to better understand the characteristics of its customers. You have been asked to perform a market basket analysis to analyze customer data to identify key associations of your customer purchases, ultimately allowing better business and strategic decision-making.

List of Tables

List of Figures

Part I: Research Question

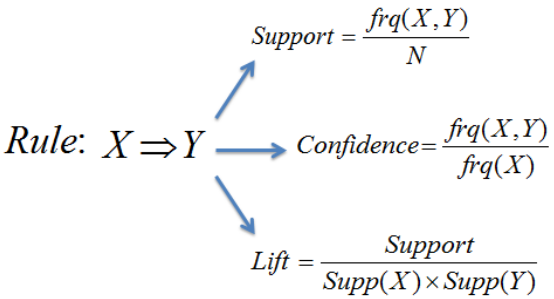
A. Describe the purpose of this data analysis by doing the following:

A1. Summarize one research question that is relevant to a real-world organizational situation captured in the selected data set and that you will answer using time series modeling techniques.

A2. Define the objectives or goals of the data analysis. Ensure that your objectives or goals are reasonable within the scope of the scenario and are represented in the available data.

Figure 1

Market Basket Analysis metrics, small visual example



Example:

Rule	Support	Confidence	Lift
$A \Rightarrow D$	2/5	2/3	10/9
$C \Rightarrow A$	2/5	2/4	5/6
$A \Rightarrow C$	2/5	2/3	5/6
$B \& C \Rightarrow D$	1/5	1/3	5/9

Notes. Figure courtesy of Susan Li (Li, 2017)

Part II: Method Justification

B. Summarize the assumptions of a time series model including stationarity and autocorrelated data.

Part III: Data Preparation

C. Summarize the data cleaning process by doing the following:

C1. Provide a line graph visualizing the realization of the time series.

C2. Describe the time step formatting of the realization, including any gaps in measurement and the length of the sequence.

C3. Evaluate the stationarity of the time series.

C4. Explain the steps used to prepare the data for analysis, including the training and test set split.

C5. Provide a copy of the cleaned dataset.

Part IV: Model Identification and Analysis

D. Analyze the time series dataset by doing the following:

D1. Report the annotated findings with visualizations of your data analysis, including the following elements:

- the presence or lack of a seasonal component
- trends
- auto correlation function
- spectral density
- the decomposed time series
- confirmation of the lack of trends in the residuals of the decomposed series

D2. Identify an autoregressive integrated moving average (ARIMA) model that takes into account the observed trend and seasonality of the time series data.

D3. Perform a forecast using the derived ARIMA model.

D4. Provide the output and calculations of the analysis you performed.

D5. Provide the code used to support the implementation of the time series model.

Part V: Data Summary and Implications

E. Summarize your findings and assumptions, including the following points:

E1. Discuss the results of your data analysis, including the following:

- the selection of an ARIMA model
- the prediction interval of the forecast
- a justification of the forecast length
- the model evaluation procedure and error metric

E2. Provide an annotated visualization of the forecast of the final model compared to the test set.

E3. Recommend a course of action based on your results.

Part VI: Reporting

F. Create your report from part E using an industry-relevant interactive development environment (e.g., a Jupyter Notebook). Include a PDF or HTML document of your executed notebook presentation.

G. List the web sources used to acquire data or segments of third-party code to support the application.

H. Acknowledge sources, using in-text citations and references, for content that is quoted, paraphrased, or summarized.

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

References

Bruce, P. C., Gedeck, P., Shmueli, G., & Patel, N. R. (2019). *Data Mining for Business Analytics Concepts, Techniques and Applications in Python.* Wiley & Sons, Incorporated, John.

Li, S. (2017, September). A Gentle Introduction on Market Basket Analysis — Association Rules. *A Gentle Introduction on Market Basket Analysis — Association Rules*. Retrieved May 24, 2022, from https://towardsdatascience.com/a-gentle-introduction-on-market-basket-analysis-association-rules-fa4b986a40ce

Maitra, S. (2021, September). Association Rule Mining using Market Basket Analysis. *Association Rule Mining using Market Basket Analysis*. Retrieved May 24, 2022, from https://towardsdatascience.com/market-basket-analysis-knowledge-discovery-in-database-simplistic-approach-dc41659e1558

Market Basket Analysis. (2015, September). *Market Basket Analysis*. Retrieved May 26, 2022, from https://sarahtianhua.wordpress.com/portfolio/market-basket-analysis/

Mazzeschi, M. (2020, June). Performing a Market Basket Analysis with Machine Learning. *Performing a Market Basket Analysis with Machine Learning*. Retrieved May 24, 2022, from https://pub.towardsai.net/performing-a-market-basket-analysis-with-machine-learning-94b5514ce1cb

Pipis, G. (2021, May). Market Basket Analysis and Association Rules from Scratch. *Market Basket Analysis and Association Rules from Scratch*. Retrieved May 24, 2022, from https://medium.com/geekculture/market-basket-analysis-and-association-rules-from-scratch-b621dd55e776

Roshan, B. (2020, December). Market Basket Analysis- Association Rule Mining with visualizations. *Market Basket Analysis- Association Rule Mining with visualizations*.

Retrieved May 26, 2022, from https://medium.com/analytics-vidhya/market-basket-analysis-association-rule-mining-with-visualizations-cda24d537019

Sivek, S. C. (2020, November). Market Basket Analysis 101: Key Concepts. *Market Basket Analysis 101: Key Concepts*. Retrieved May 26, 2022, from https://towardsdatascience.com/market-basket-analysis-101-key-concepts-1ddc6876cd00

Tang, L., Li, J., Du, H., Li, L., Wu, J., & Wang, S. (2022, February 28). Big Data in Forecasting Research: A Literature Review. *Big Data Research, 27*, 100289. doi:10.1016/j.bdr.2021.100289

Yıldırım, S. (2020, October). Market Basket Analysis with Pandas. *Market Basket Analysis with Pandas*. Retrieved May 24, 2022, from https://towardsdatascience.com/market-basket-analysis-with-pandas-246fb8ee10a5