

BMKT 491: Special Topics, Advanced Marketing Analytics Assignment

- Omar Shehryar (Marketing Option, JABS College of Business and Entrepreneurship)

Assignment Notes: The assignment is designed as a test, either mid-term, or final, in a course titled “Advanced Marketing Analytics.” Currently this course is not hard-numbered but has been offered twice as a special topics course with a place-holder course number, BMKT 491. The course is focused on the use of machine learning for marketing decision making. Students are introduced to machine learning techniques and their applications in marketing. Students are provided code during the semester. Students may access this code during the test, but the data that will be used will not be known to students. Hence, data wrangling will be required. Students may use AI to change the structure of the data to fit it for use with the algorithm that is required.

License: This assignment is licensed under [CC-BY-NC](#)

Overview: In this assignment, students will take on two machine learning tasks. The first task is a simpler and more widely used machine learning task of data reduction. Students will employ k-means clustering algorithm to find useful patterns in the data to determine actionable marketing segments. The second task requires students to perform association-rules mining on retail data. This task is more substantial and requires both data wrangling, and application of the Apriori algorithm.

Purpose: This assignment is designed to help students think critically about data shape, algorithm selection appropriate for the task, useful and un-useful data, and making marketing decisions based on big data that cannot be analyzed with just spreadsheets. Thus, students are expected to perform the same tasks that a contemporary data scientist performs. Students will also become familiar with the use of ChatGPT as a coding assistant to tackle the data wrangling tasks.

Duration: Time allotted is two class periods of 75 minutes each.

Learning Objectives

By the end of this assignment, students will be able to:

1. Import and Shape data to fit the requirements of the task
2. Select the appropriate algorithm(s) and execute appropriate code to run the algorithm(s)
3. Visualize important results obtained from R programming language output in R studio
4. Use AI assistance to complete their assigned tasks

Generative AI Tool for Assignment

For more in-depth information on this AI tool, see attached model card.

- Tool Name: ChatGPT
- Tool Purpose: ChatGPT is primarily a large language model and has myriad uses. However, for the tasks at hand, the purpose of using the tool is to provide students an assistant that

can help them with data wrangling tasks, resolve issues with importing data, and guidance with errors in coding

- Developer/Version: Open AI/ChatGPT-4o
- Known Limitations/Biases: ChatGPT primarily uses English Language texts for model training. It has known gender biases as well c.f. Babaei, G., Banks, D., Bosone, C., Giudici, P., & Shan, Y. (2024). Is ChatGPT More Biased Than You? Harvard Data Science Review, 6(3). <https://doi.org/10.1162/99608f92.2781452d>

Assignment Structure

1. Perform k-means clustering on the provided data file titled “test.xls.”
 - a. Load the data and shape appropriately for analysis (30 points).
 - b. What is the optimal number of clusters? (5 points)
 - c. Label the 2-dimensional cluster solution and interpret the key findings from the cluster solution. (15 points).
2. Perform Market Basket Analysis on the E-Commerce Market Basket Data file.
 - a. Load the data, and shape it appropriately for market basket analysis (30 points)
 - b. What is the total number of rules found? (5 points)
 - c. List and explain the top 3 association rules including the following (15 points):
 - i. State the rules in simple language.
 - ii. Explain the support, confidence, and lift numbers and what they mean for the rule, including the names of the products involved.

Submission Guidelines

- Answer the questions in a PDF document. This document will also contain any prompts you used to get assistance from ChatGPT.
- Separately attach RStudio visualization file for 1c. and RStudio output file for 2c.
- Length: The PDF file with answers needs to be no longer than one page. The RStudio visualization for the 2-dimensional cluster solution is an additional page. Lastly, the top-association rules output is an additional page.
- IMPORTANT: Save your code file and upload it as proof of your work.
- Citation Guidelines: If you used ChatGPT for assistance, please cite it in the following format:
 - OpenAI. (2023). ChatGPT-4o (Oct 08, 2024 version) [Large language model]. <https://chat.openai.com/chat>

Grading Rubric

Points are awarded based on completion of each individual element of the assignment, as noted after each element.

Additional Resources

Use of code including algorithms and their use, and code chunks involved in data wrangling i.e., tidyverse operations, is permitted and encouraged.

Notes on Model Card Intended Use

Model card is intended for classroom educational use only. Model card may be used by any educator whose intended use is classroom instruction.