**Topic:** Understanding consumer buying behaviour using machine learning

Student name:

Student ID:

**Introduction**

Proper marketing strategies are essential for organizations to attract more customers and improve sales and their profits (Osterwalder et al., 2015). To achieve this, companies have to understand their customer or consumer behaviours. Consumer behaviour are decisions and actions that characterize customers who purchase goods and services from stores, either on-store or online. Studying what drives people to purchase items that satisfy them is important for organizations as they can anticipate what products that their customers want to buy. The marketing department can then design their marketing campaigns based on the consumer behaviour findings. It will be vital for a company to understand the success of its products and the new products to be launched as well. There are high chances of product failure if a company does not understand the reaction of their customers towards their products (Bai et al., 2015). Consumer behaviour changes due to factors such as trends, technology, fashion, living style, and disposable income, among other options. Understanding consumer behaviour helps organizations to differentiate their consumers into respective demographics and serve them based on their unique needs. On top of attracting new customers, it also helps in retaining the existing ones. An organization can create relevant marketing programs based on the needs or behaviours of the target consumers. Customer service is improved as an organization provides services based on how they have understood their customer purchase behaviours.

The process of understanding consumer buying behaviour entails a comprehensive analysis of user-generated content such as ratings, reviews, comments, or purchase quantities (Sethna et al., 2017). The analysis results can help the organization better understand its clients and enable the management to modify their products or services based on different customer purchase trends and needs. The organization can also use the insights to predict future purchase intentions towards specific items or services. The insights generated from the analysis can help the management to track the attitudes associated with their clients. The insights can also be used to create recommender systems that offer customised buying experience for the customers (Sethna et al., 2017). All these benefits are intended to tailor products and services based on client needs to boost sales and improve profits. This report will entail implementing a recommender system based on the analysis of the customer purchase behaviour for an online store. Machine learning algorithms will be used to build a recommender system that proposes for a customer what to purchase based on the items the customer needs has selected for purchase.

**Problem statement**

Businesses have to meet consumer needs in order to realize a stable growth. This implies that organizations have to look for ways to adapt to the constantly emerging trends in the market. They have to understand the consumer needs and demands by experimenting strategies that work and engage their consumers. Marketers need to understand consumer purchase behaviour and decisions in order to optimize their engagement with their clients. One of the toughest problems that marketers face is finding key product or item pairs that customers frequently purchase together. This is where the concept of recommender system comes in to filter items overload based on the interests, preferences, and observed behaviour of the user. The system beneficial to both the user and the company selling goods and services. The costs for finding and selecting products in an online shopping experience are greatly reduced.

**Proposed solution**

This report emphasizes on the implementation of a recommender system for an online store that will be designed using appropriate machine learning algorithms to help predict preferable items to purchase based on what the user has selected to buy.The system will anticipate the needs of the customers for the online store as products will get personalized based on the purchase patterns of the customer. The next item to purchase is suggested by the system. This maximizes the probability of purchase. The system will improve the user experience and will make it easier for the customers to find the products that interest them. This in turn, improves customer satisfaction and boost sales and revenues. The associated company builds a stronger customer loyalty from the satisfaction they get while purchasing products. Churn rates are drastically reduced, which reduces the costs the associated company incurs while acquiring new clients.

**Aim:** The aim of the project is to understand the customer purchase behaviour using data mining techniques and building a recommender system that will propose the next product to add to the basket.

**Objectives**

The objectives of the report are:

* Investigate what constitutes consumer purchase behaviour
* Research on how to capitalize on consumer purchase behaviour
* Investigate how data mining techniques can be used to understand consumer buy behaviour
* Study the types of recommendations offered by data mining or machine learning tools
* Design a recommender system with appropriate machine learning algorithms for an online store that runs on web

**Methodology**

I intend to build a data mining system that understands customer purchase intentions and recommends the next product to purchase based on what is the product in the cart. I will use appropriate machine learning algorithms, specifically the ones based on association rules. The algorithm used will calculate the item frequencies and group the frequently purchased items together. A recommender system will be designed that studies the purchase frequencies of the items that customers buy. The system focuses on two entities – the customer and the product.

The algorithms applied in building a recommender system are memory-based and are categorized under the collaborative filtering category. They work on the user-item matrix to give a prognosis. The recommendations performed by the system are classified as either association rule-based or most-frequent item recommendations. This research focuses on association rule-based recommendation. The association between the items is calculated using concepts like support, confidence, lift, and conviction. Support is essentially is an indication that shows how often an item-set appears in the dataset. Confidence indicates how often a rule has been found to be true. Lift is the correlation of support for an item set for self-sufficient values. Conviction is the ratio of the occurrence rate when a rule makes a wrong prediction or the expected frequency that one item occurs without the other.

The implementation will analyse transactional data to deduce buying patterns from the UCI Machine Learning Repository, [link](https://archive.ics.uci.edu/ml/datasets/online+retail). It contains the transaction history of an online store over 1 year. It contains information about customer purchases such as item descriptions, invoice, unit price, and country of the customer’s origin. It is from the customer buying patterns that association rules will be created. These rules will be used to make recommendations based on what the customer has purchased. Among the many data mining algorithms that can be used for this purpose, Apriori algorithm will be selected. It is one of the most classical algorithms in data mining that uses association rules.

The image below summarizes the steps to be followed in the data analysis before the recommender system is build.

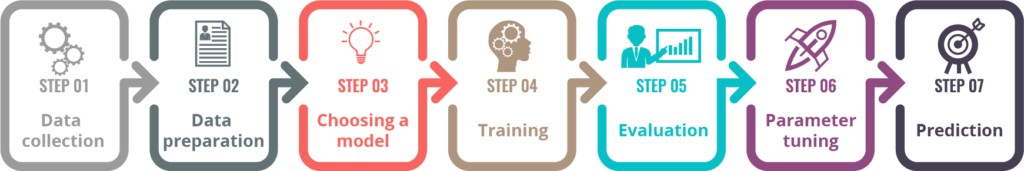


Figure 1: Steps in the data analysis

The first step is sourcing appropriate data that will train the model to be built. This is followed by data preparation. When the data is not clean, the developed models are not robust and are prone to give inaccurate results. It entails removing noisy data from the dataset and dealing with missing values. The next step will be choosing the right model from the many machine learning models available depending on the task. Here, association rule-based algorithms will be used to determine frequently purchased items. The model will be trained and evaluated against appropriate metrics. Parameter tuning will also be performed to optimize the performance of the final model. The final model will then be used to create a recommender system that will predict items that a customer should purchase based on what he/she has selected.

**Project plan**

|  |  |  |  |
| --- | --- | --- | --- |
| **Date** | **Task** | **Duration (days)** | **Progress** |
| 03/3/2022 | Topic proposal | 3 |  |
| 10/3/2022 | Chapter 1: Introduction | 5 |  |
| 20/3/2022 | Chapter 2: Literature Review | 6 |  |
| 30/3/2022 | Chapter 3: Research methodology | 5 |  |
| 10/4/2022 | Chapter 4: Implementation | 10 |  |
| 20/4/2022 | Chapter 5: Results and findings | 7 |  |
| 30/4/2022 | Chapter 6: Conclusion and recommendations | 5 |  |

**References**

Bai, Y., Yao, Z., & Dou, Y.-F. (2015). Effect of social commerce factors on user purchase behavior: An empirical investigation from renren. com. *International Journal of Information Management*, *35*(5), 538-550.

Osterwalder, A., Pigneur, Y., Bernarda, G., & Smith, A. (2015). *Value proposition design: How to create products and services customers want* (Vol. 2). John Wiley & Sons.

Sethna, B. N., Hazari, S., & Bergiel, B. (2017). Influence of user generated content in online shopping: impact of gender on purchase behaviour, trust, and intention to purchase. *International Journal of Electronic Marketing and Retailing*, *8*(4), 344-371.