

# Applying Clustering Technique and Association Rule to Analyze Laptop Usage Behavior of Students\*

Nguyen Van Chuc\*\* [0000-0003-2618-5659] and Nguyen Ha Nhu Ngoc

University of Economics, The University of Danang, Vietnam  
chuc.nv@due.edu.vn, ngocnhn43k08@due.udn.vn

**Abstract.** A laptop is a typical technological product with high mobility qualities that allows everyone to learn and work from anywhere. These days, laptops are in high demand, particularly among students. There are numerous competing brands on the market with full lines, varieties, configurations, and prices ranging from inexpensive to high-end, making it difficult for customers to buy. Analyzing the behavior of students using laptops to discover trends and factors influencing their decision to buy a laptop and thus assisting them in making the best choice when choosing. It is also extremely beneficial for laptop distributors and merchants because it helps them to reach out to a larger number of potential customers. The article focuses on applying clustering techniques and association rules in data mining to analyze the laptop usage behavior of students. Some solutions are provided based on the acquired results to assist organizations in understanding customer characteristics and making better business decisions.

**Keywords:** Data Mining, Behavior Analysis, Clustering, Association Rule.

## 1 Introduction

The Information Age is rapidly and strongly evolving, resulting in the birth of a slew of extremely modern and intelligent electronic devices. It is impossible to discuss smart devices without mentioning laptops. Recently, as a result of the global pandemic of Covid-19, there has been an increase in the use of mobile devices such as laptops for communication, distance learning, and knowledge learning based on Google's platform. There are currently too many laptop lines on the market, making it difficult for users to choose the brand, function, and price that are reasonable and best suited to their personal use requirements. It's also a common question among students. Faced with this reality, distributors and retailers must understand customer psychology, needs, and preferences in order to develop effective business policies, advertising, and marketing

---

\* Copyright © by the paper's authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0). In: N. D. Vo, O.-J. Lee, K.-H. N. Bui, H. G. Lim, H.-J. Jeon, P.-M. Nguyen, B. Q. Tuyen, J.-T. Kim, J. J. Jung, T. A. Vo (eds.): Proceedings of the 2nd International Conference on Human-centered Artificial Intelligence (Computing4Human 2021), Da Nang, Viet Nam, 28-October-2021, published at <http://ceur-ws.org>

\*\* Corresponding author.

strategies to increase market share and attract customers, particularly students-a potential customer source for this item.

## 2 An overview of data clustering techniques and association rules

### 2.1 An introduction to clustering techniques

Data clustering is the process of grouping given objects into clusters so that objects in the same cluster are as similar as possible and objects in different clusters are as different as possible. The goal of clustering is to determine the inner groupings of data. There are numerous clustering techniques available, including partition clustering, hierarchical clustering, density-based clustering, and so on [1].

### 2.2 An introduction to association rules

The goal of Association Rules (AR) in data mining is to find relationships between objects in large amounts of data. The fundamental of the AR is summarized [2]. Given the transaction database  $T$  contains the set of transactions  $t_1, t_2, \dots, t_n$ .

$T = \{t_1, t_2, \dots, t_n\}$ . Each transaction ( $t_i$ ) is made up of a set of objects  $I$  (itemset).

$I = \{i_1, i_2, \dots, i_m\}$ . A  $k$ -itemset is an itemset made up of  $k$  items.

The purpose of AR is to discover associations (correlations) between items. These association rules take the form of  $X \rightarrow Y, X \cap Y = \emptyset$  (1)

Where  $X$ : antecedents;  $Y$ : consequents

Support and confidence are two crucial criteria in evaluating association rules. The formula for calculating the support and confidence of the association rule  $X \rightarrow Y$  [2, 3]:

$$\text{support}(X \rightarrow Y) = P(X \cup Y) = n(X \cup Y)/N \quad (2)$$

$$\text{confidence}(X \rightarrow Y) = P(Y|X) = n(X \cup Y)/n(X) \quad (3)$$

Where  $n(X)$ : Number of transactions containing  $X$ ;  $N$ : Total number of transactions.

AR with support and confidence greater than or equal to the minimum support (min sup) and minimum confidence (min conf) are referred to as strong rules [2, 3].

## 3 Using clustering technique and association rule to analyze student laptop usage behavior

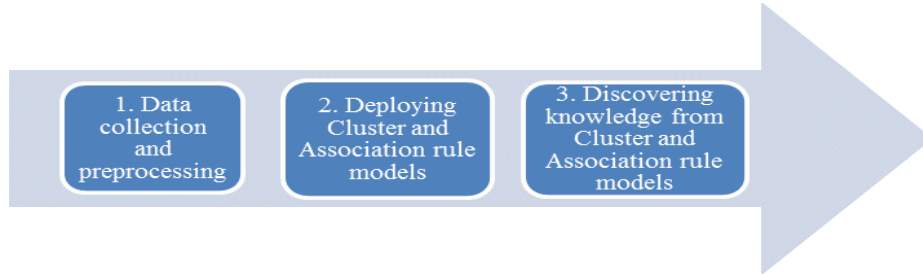
### 3.1 Problem description

Clustering techniques and association rules are used to analyze the laptop usage behavior of students.

**Input:** information about students, laptops and factors influencing purchase.

**Output:** provide the characteristics, behavior of using laptops across majors, predict usability, and the relationship between factors when students decide to buy a laptop.

### 3.2 Implementation scenario for student laptop usage behavior analysis system



**Fig. 1.** Implementing a student laptop behavior analysis system

**Step 1:** Data collection and preprocessing. From March to May 2021, 1100 samples were collected through an online questionnaire survey of students from University of Economics - The University of Danang. After data preprocessing, Table 1 shows the data structure.

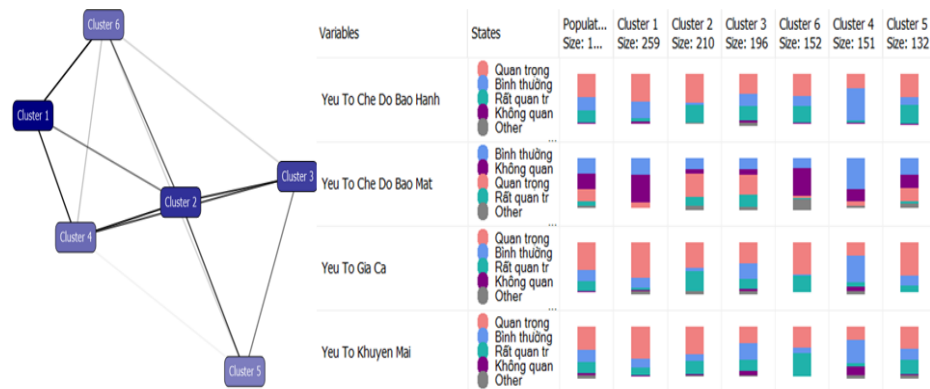
**Step 2:** Deploying cluster and association rule models. The model was created using Microsoft Business Intelligence Development Studio (BIDS) data mining tool. The discovered knowledge is very intuitive, easy to understand, and simple to apply [4, 5].

**Step 3:** The knowledge discovered from the Cluster and Association rule models used to analyze laptop usage behaviour of students. (Fig.1)

**Table 1.** Data description

No.	Attribute name	Data types	Value domain	Meaning
1	ID	Interval	1 - 1100	Individual survey order
2	GioiTinh	Nominal	Male, Female	Gender of the students
3	QueQuan	Nominal	Da Nang, Hue...	The student's hometown
4	NamHoc	Nominal	1 <sup>st</sup> , 2 <sup>nd</sup> year	Year students are studying
5	Khoa	Nominal	Banking...	Faculty
6	Nganh	Nominal	Accounting...	Major
7	NgheNghiepGD	Nominal	Farmers...	Parents' occupations
8	ChiTieuHangThang	Nominal	1.5 million VND	Monthly budget for a student
9	ThuongHieu	Nominal	Asus, Dell...	Brand of laptop
10	ThoiGianMua	Nominal	Under 6 months	When purchasing a laptop
11	Gia	Nominal	15 million VND	Laptop pricing
12	MucDichSuDung	Nominal	Studying, ...	Purpose of buying a laptop
13	MucDoHaiLong	Interval	1 → 5	Level of satisfaction
14	ThongTinMua	Nominal	Websites...	Sources of information
15	NoiMua	Nominal	FPT Shop ...	Laptop stores
16	YeuToThuongHieu	Nominal	Very important...	Factor of Brand

17	YeuToCauHinh	Nominal	Important...	Factor of Configuration
18	YeuToTocDoXuLy	Nominal	Normal...	Factor of Processing Speed
19	YeuToGiaCa	Nominal	Unimportant...	Factor of Price
20	YeuToKieuDang	Nominal	Normal...	Factor of Style
21	YeuToUyTin	Nominal	Very important...	Factor of Retailer Reputation
22	YeuToBaoMat	Nominal	Very important...	Factor of Confidential Mode
23	YeuToBaoHanh	Nominal	Very important...	Factor of Warranty
24	YeuToKhuyenMai	Nominal	Very important...	Factor of Promotion



**Fig. 2.** The results of data clustering

According to the model's results (Fig.2), Table 2 shows six clusters:

**Table 2.** Cluster characteristics

Cluster	Size	Cluster characteristics	
Cluster 1	259 (23.5%)	Female in 3 <sup>rd</sup> year, parents are farmers, the cost between 10 and 15 million VND. Seeking information from relatives and friends, expert → Dell, Asus.	- Configuration, Processing speed: Normal; Brand, Warranty, Promotion, Price, Retailer reputation, Style: Important; Security: Unimportant.
Cluster 2	210 (19.1%)	Female in 4 <sup>th</sup> year, parents are farmers, and the cost between 10 and 15 million, according to information obtained from relatives and friends → Dell.	- Processing speed, Style: Normal; Brand, Configuration, Warranty, Security, Promotion, Retailer reputation: Important; Price: Very important.
Cluster 3	196 (17.8%)	Female in 2 <sup>nd</sup> , 3 <sup>rd</sup> years, parents are government employees, self-employed. It cost over 25 million VND to seek information from family and sales-people → Apple.	- Promotion: Normal; Price, Configuration, Processing speed, Warranty, Security: Important; Brand, Retailer reputation, Style: Very important.
Cluster 4	151 (13.7%)	Male in 2 <sup>nd</sup> and 3 <sup>rd</sup> years, parents are farmers, and the cost	- Brand, Warranty, Security, Promotion, Retailer reputation: Normal;

Cluster 5	132 (12.0%)	from 10 to 15 million VND, seeking purchasing information salespeople → Dell and Asus. Male in 3 <sup>rd</sup> year, parents are farmers and self-employed. The cost between 15 and 20 million VND. They seek purchasing information from salespeople and expert guidance → Dell and HP.	Price, Configuration, Processing speed: Important; Style: Unimportant.
Cluster 6	152 (13.8%)	Female in 2 <sup>nd</sup> year, parents are farmers with laptop prices under 10 million VND, referring to buying information from family, friends, and salespeople → Asus and Acer.	- Security: Normal; Promotion, Brand, Warranty, Price, Retailer reputation: Important; Style: Unimportant; Configuration, Processing speed: Very important.



Fig. 3. The results of association rule model.

Here are some association rules (Fig.3):

R1: With a price of 25 million VND, second-year Foreign Trade students in Da Nang are interested in brand design and security issues when purchasing a laptop that primarily uses the Apple brand, with 100 percent confidence.

R2: Students of Commerce and Management Information Systems (MIS) from Quang Nam, Ha Tinh, are interested in warranty, configuration, and brand factors when purchasing a laptop with a price range of 10 to 15 million VND. Dell is the most commonly used brand, with 75 percent confidence.

R3: With a price range of 10 to 15 million VND, third-year Accounting students in Dak Lak, who are concerned with price, promotion, and warranty when purchasing a laptop, primarily use the Asus brand, with 63 percent confidence.

R4: With a price of less than 10 million VND, Accounting and Commerce students in Quang Nam and Hue, are interested in the price, promotion, and warranty factors when purchasing laptops that primarily use the Asus and Acer brands, with 60 percent confidence.

R5: With a price range of 15 to 20 million VND, a male student majoring in MIS whose parents are freelancer, is interested in configuration factors, processing speed, and warranty when purchasing HP laptops, with 53 percent confidence.

## **4 Proposing data-driven marketing and CRM solutions**

### **4.1 Describing product feature**

The clustering results show that Dell and Asus are two brands that students frequently purchase, with prices ranging from 10 to 15 million VND. To minimize shortages, distributors should concentrate on these two important brands.

The findings of the association rule highlight some aspects of various majors:

#### *About the brand*

The majority of Foreign Trade students utilize Apple products; Students of MIS like the HP brands, which cost 15-20 million VND; Students of Accounting and Business choose Asus and Acer brands with prices under 10 million VND.

#### *About the factors*

Foreign Trade students are interested in design and security factors; MIS students are interested in configuration and processing speed factors; Accounting, Business and Commerce students are interested in price and promotion factors.

### **4.2 Sources of information for purchasing a laptop**

Improve image promotion, coverage and word-of-mouth marketing: Use the credibility of celebrities and loyal customer groups.

Enhance product quality and the company's reputation; develop after-sales customer care plans to encourage customers.

- Organize the hiring and training of professional staff who are familiar with laptops and have a professional service style. Training and coaching staff to always smile at customers in all situations.

- Create a separate customer service department to handle customer feedback. All requests for purchasing, selling and delivering services are handled in a timely manner. Companies can also set up separate phone lines to handle customer inquiries.

### **4.3 Programs for advertising and marketing**

Companies must improve their communication on social networking sites, as well as websites that combine programs, fairs, and exhibitions held at schools. Updating information on e-commerce channels to disseminate knowledge and general information about product features.

According to the clustering results, students are subjects with a limited budget who place a high value on price and promotion. As a result, distribution and retail companies should:

- Implement a marketing strategy suitable to students of different majors according to their behaviours and characteristics.
- Offer optimal payment terms, such as implementing an installment policy, purchasing first and paying later with a 0% interest rate.
- Put product discounts, promotions, and giveaways into action. Promotions must be diverse and of high quality.
- There are policies in place to allow customers to return goods and receive refunds in certain circumstances, such as when laptop malfunctions.

## 5 Conclusions and future work

The article studied the theory of clustering techniques and association rules for applying these techniques to build data mining models to analyze the laptop usage behavior of students. The results of the clustering technique analysis have clarified the outstanding features of groups of students with similar characteristics; the association rules discovered from the data help to understand the relationship and influence of factors influencing students' choice to buy laptops. The knowledge extracted from the models assists laptop distributors and retailers in understanding the trends and characteristics of students, allowing them to develop effective business strategies. More data will be gathered in the coming months from a variety of sources, including not only students from University of Economics-The University of Danang, but also extensive research with students from all over Danang City, making the data more complete, in order to improve the model and increase the efficiency of analysis and prediction.

## References

1. Chuc, N.V. and Giang, D.T.: Applying clustering technique and association rule to exploit data of customers using hotel services. Journal of science and technology (JST-UD), **12**(97), 109-112 (2015).  
[http://tapchikhn.udn.vn/volume/135/Ung\\_dung\\_ky\\_thuat\\_phan\\_cum\\_va\\_luat\\_ket\\_hop\\_khai\\_pha\\_du\\_lieu\\_khach\\_hang\\_su\\_dung\\_dich\\_vu\\_khach\\_san-8501/8501](http://tapchikhn.udn.vn/volume/135/Ung_dung_ky_thuat_phan_cum_va_luat_ket_hop_khai_pha_du_lieu_khach_hang_su_dung_dich_vu_khach_san-8501/8501)
2. Tan, P.N, Steinbach M., Karpatne, A., Kumar V.: Introduction to data mining, 2<sup>nd</sup> Edition, Pearson (2018).
3. Witten, L.H, Frank, E., Hall, M.A., Pal, C.J.: Data mining: Practical machine learning tools and techniques, 4<sup>th</sup> Edition, Morgan Kaufmann (2016).
4. SQL Server Analysis Services (SSAS). <https://docs.microsoft.com/en-us/analysis-services/data-mining/data-mining-ssas?view=asallproducts-allversions>.
5. Data Mining and Business Intelligence solution. <http://bis.net.vn/forums/36.aspx>