ORIGINAL RESEARCH PAPER

Marketing

KEY WORDS: online shopping, website comparison, paired similarity index, e-commerce

ANALYSIS OF SELECT INDIAN E-COMMERCE WEBSITES USING PAIRED SIMILARITY INDEX

Debanuj Khound Research Scholar, Centre for Management Studies (CMS), Dibrugarh University

Dr. Bipasha Chetiya Barua* Assistant Professor, Department of Commerce, Dibrugarh University*Corresponding Author

ABSTRACT

Online shopping industry is one of the fastest growing industries which has expanded rapidly due to the ever increasing access to internet and rising disposable income levels. It is observed that new age consumers are indulging in impulsive purchase of products in the online shopping website. The information content, features, web design etc. in the websites plays a significant role in drawing shoppers. In this paper the information and features provided in some popular e-commerce websites are analysed using binary representation and developing a similarity index. A Paired Similarity Index (PSI) is created to measure the similarity of information and features between websites through pairs of observations. The study is comprised of selected e-commerce websites, operating in India across a different product categories. This comparison will enable to rank websites based on the information and features provided.

Introduction

Online shopping is a major part of the overall electronic commerce, or e-commerce, industry which consists of all the buying and selling of goods and services over electronic systems such as the internet and other computer networks by households, businesses and other agencies. The Indian market is considered as one of the most lucrative owing to the increasing purchasing power of customers and a favourable young, tech savvy demography. This has led to immense competition among marketers to grab the larger part of the share of wallet. Today, the Indian consumer have access to all global brands and it has been made easier with the growth of e-commerce industry. Online shopping is a major part of the overall electronic commerce, or e-commerce, industry which consists of all the buying and selling of goods and services over electronic systems such as the internet and other computer networks by households, businesses and other agencies. The e-commerce market in India is expected to reach US$ 7 trillion in 2023, growing at a CAGR of 20% in the last five years (Ankur, 2021).

A typical online shopper spends about 30 seconds viewing a website before they decide to click away or view merchandise on the website (Brohan, 1999). So the information and features provided in an e-commerce website is very crucial to draw online shoppers in order to make them purchase products.

For the purpose of this study 12 popular Indian e-commerce websites are selected, who are engaged in selling of the following product categories (at least one category):

a) Personal Gadgets (e.g. Mobile Phones, Tablets, Smart Watches etc.)
b) Home Appliances (e.g. Television, Refrigerator, Air Conditioner etc.)
c) Furniture and Home Décor (e.g. Household furniture, Decorative items, Lights etc.)
d) Precious Jewellery (e.g. Gold, Diamond, Platinum jewellery items)

The information contained and features provided by these 12 websites were classified into four broad categories and again categorized under some sub-categories. If a particular information or feature is provided by the e-commerce website, it is coded as '1' and otherwise coded as '0' denoting absence. In this paper independence between the categories is assumed.

Objectives of the Study

This research paper propose to study the following

objectives:

a) To identify the similarity of e-commerce websites of India by using the Paired Similarity Index (PSI)
b) To prepare a ranking of the e-commerce websites on the basis of the information and features provided by them

Methodology

For the purpose of this study, the following popular e-commerce website are considered: Amazon, Flipkart, Snapdeal, Shopclues, Myntra, Koovs, Ajio, TataCliq, UrbanLadder, PepperFry, Bluestone, Caratlane. The information and features provided by the above mentioned e-commerce websites are broadly categorized under four heads, namely:

a) Product Specifications
b) Payment Options
c) Customer Services
d) Web Design

Under these four categories a total of 41 attributes are considered for comparison among the 12 e-commerce websites, details of which is provided in Appendix-A. The relevant data is collected from the mentioned e-commerce websites in the month of September, 2021.

PSI for Binary Data

The research work of (Erlish, 2002) proposed a procedure for data mining by using binary representation for determining Paired Similarity Index (PSI) between any two entities. The information content and features provided by the e-commerce websites are classified into four broad categories and under each category certain attributes are considered. For representing the attributes, a binary vector is used for denoting the presence (1) and absence (0).

As proposed by (Erlish, 2002), the measure of similarity can be explained as follows:

Suppose that for each website 'i' (i= 1, 2, 3 ...., n) we have 'm' categories. For each category 'j' (j= 1, 2, ..., m) we have pj attributes. The value of pj is called as the domain size of the jth category. They define the binary representation vector of length, $p = \sum_{j=1}^{m} p_j$ (the length of the domain category vector), for each website 'i' (i= 1, 2, 3 ...., n) in the following way: X_{ijk} = 1, if the information about the kth attribute belonging to the jth category is present in ith website, absence is = 0.

Where I = 1,2,3 ….,n, then j = 1,2,3 ….,m and k = 1,2,3 ….P_j

The mutually exclusivity property for each category over its domain was assumed. Using binary representation (Erlish, 2002) defined a Paired Similarity Index (PSI) as follows:

$$PSI = PSI(i_1, i_2) = \frac{sa(i_1, i_2)}{m} \quad (1)$$

Where $m = \sum_{j=1}^{m} \sum_{k=1}^{p_j} x_{ijk}$ and $sa(i_1, i_2) = \sum_{j=1}^{m} \sum_{k=1}^{p_j} x_{i_1 j_k} = \sum_{j=1}^{m} \sum_{k=1}^{p_j} x_{i_1 j_k}$

Now for each category j, if a website can attain maximum possible of its Pj domain values (I,e when the mutually exclusivity property doesn't satisfy for each category over its domain) then the range of paired similarity index (PSI) given by (Erlish, 2002) is greater than one (i,e PSI > 1). If the value of PSI is greater than 1, then it is difficult to determine the similarity measure between any two e-commerce websites. Thereby, an absolute similarity between two websites cannot be designated in case of binary representation using (1).

A new similarity index, as ratio between the number of similar attribute values of any two websites and the length of the domain attribute vector to overcome the above mentioned difficulties for non-mutually exclusive cases was designed by (Saikia & Bhattacharjee, 2009) The redefined PSI developed is given below:

$$PSI = \frac{sa(i_1, i_2)}{m} \quad (2)$$

Where $p = \sum_{j=1}^{m} p_j$ and $sa(i_1, i_2) = \sum_{j=1}^{m} \sum_{k=1}^{p_j} x_{i_1 j_k} = \sum_{j=1}^{m} \sum_{k=1}^{p_j} x_{i_1 j_k}$

Now the similarity index range will be $0 \leq PSI \leq 1$, where PSI = 1 indicates absolute similarity and PSI = 0 indicates absolute difference between any two e-commerce website.

Example

For calculating PSI between two e-commerce websites, say Flipkart (i_2) and Snapdeal (i_3) as given in Appendix – B, using formula (2)

$$PSI = \frac{sa(i_2, i_3)}{m}$$

In this case, m = 41, and $sa(i_2, i_3) = \sum_{j=1}^{4} \sum_{k=1}^{41} x_{i_2 j_k} = 23$

so PSI = 23/41 = 0.561

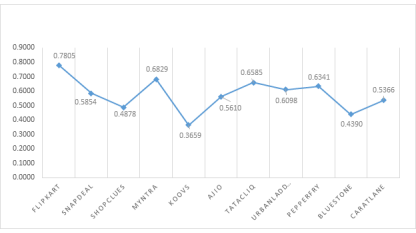
Here the PSI of 0.561 indicates a very low similarity between the websites of Flipkart and Snapdeal.

In the same way the PSI for the 12 e-commerce websites are calculated and the values of the corresponding PSI scores are represented in a matrix given in Appendix – C.

Graphical Representation

On the basis of overall information content and features provided, Amazon ranks the first (Appendix – D). In comparison to Amazon, the PSI values of the other 11 e-commerce websites are represented in the graphical manner below:

Graph 1: Paired Similarity Index (PSI)



Findings

From the PSI Matrix given in Appendix – C and the Websites ranking done in Appendix - D, the following findings can be derived:

- The value of PSI between the websites of Amazon and Flipkart is 0.780 which is the highest, signifying a very high degree of similarity between both the websites.
- Followed by that the second highest PSI score of 0.683 is noticed Amazon-Myntra and Flipkart-Myntra.
- The lowest PSI value is 0.293, signifying a very low degree of similarity is observed between the websites of Koovs-Shopclues and Bluestone-Shopclues.
- With an overall score of 8.54, the website of Amazon occupies the Rank 1. It can be considered as the most informative and feature rich website among the 11 websites compared
- While the website of Koovs is ranked the lowest with a score of 3.90, as it is providing very less information and features to its shoppers

Conclusion

This comparative study between the popular Indian e-commerce websites provides an insight about the similarity of information content and the popular features provided by the sellers. The e-commerce websites are very dynamic in nature, so the findings may change at different time periods. A major limitation of the study is that, all the websites deals with a very large variety of products serving millions of customers across the country, so some of the features provided are dependent upon the value of the goods and the location of the customer. In future similar comparative study can be conducted for the Mobile Apps of the websites.

Appendix - A

1. PRODUCT SPECIFICATIONS
Product Description
Seller Info
Product Images
Product Videos
Product Reviews
Product Ratings
Product Comparison
Product FAQs
Product Customization

2. PAYMENT OPTIONS

Discount Vouchers/Codes
EMI Options
Zero Cost EMIs
Gift Vouchers/Cards
Partnership with Banks
E-Wallet Payments
Paybacks / Reward Points
Cancellation Fee
Cash on Delivery

3. CUSTOMER SERVICES

Postal Code Verify
Order Tracking
Delivery Customization
Delivery Charges
Return Policy
Return Charges
Call Support
Live Chat
Email Support
Membership Program
Newsletters
Gift Packings
Special Message
Installation Support
Installation Charges

3.WEB DESIGN

Today's Deal/Offer
Previously Viewed Products
Similar Product Suggestions
Frequently Bought Together Suggestions
Season Sale
Mobile App
Trial without Order
Customer Account

Appendix – B

Table 1: Product Specification Comparison

	PRODUCT SPECIFICATIONS								
Websites	Product Descriptions	Seller Info	Product Images	Product Videos	Product Reviews	Product Ratings	Product Comparison	Product FAQs	Product Customization
Amazon	1	1	1	1	1	1	1	1	0
Flipkart	1	1	1	1	1	1	1	1	0
Snapdeal	1	0	1	0	1	1	0	1	0
Shopclues	1	0	1	0	1	1	0	0	0
Myntra	1	1	1	1	1	1	0	1	0
Koovs	1	1	1	0	0	0	0	1	0
Ajio	1	1	1	0	0	0	0	1	0
TataCliq	1	1	1	0	1	1	0	0	0
UrbanLadder	1	1	1	0	1	0	1	1	1
PepperFry	1	1	1	0	0	0	0	1	1
BlueStone	1	1	1	1	0	0	0	1	1
Caratlane	1	1	1	1	1	1	0	1	1

Table 2: Payment Options Comparison

	PAYMENT OPTIONS								
Websites	Discount Vouchers/ Code	EMI Options	Zero Cost EMIs	Gift Vouchers/ Cards	Partnership with Banks	E-Wallet Payments	Pay backs /Reward Points	Cancellation Fee	Cash on Delivery
Amazon	1	1	1	1	1	1	1	0	1
Flipkart	1	1	1	1	1	1	1	0	1
Snapdeal	1	1	1	1	0	1	1	0	1
Shopclues	1	0	0	0	0	1	1	0	1
Myntra	1	1	0	1	1	1	1	0	1
Koovs	1	0	0	0	0	1	0	0	0
Ajio	1	0	0	1	1	1	1	0	1
TataCliq	1	1	1	1	1	1	1	0	1
UrbanLadder	1	1	1	1	0	1	1	1	0
PepperFry	1	1	1	1	1	1	1	1	0
BlueStone	1	0	0	1	0	1	1	0	0
Caratlane	1	0	0	1	0	1	1	0	0

Table 3: Customer Services Comparison

	CUSTOMER SERVICES														
Websites	Postal Code Verify	Order Tracking	Delivery Customization	Delivery Charges	Return Policy	Return Charges	Call Support	Live Chat	Email Support	Membership Program	Newsletters	Gift Packing	Special Message	Installation Support	Installation Charges
Amazon	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1
Flipkart	1	1	0	0	1	0	1	0	1	1	1	0	0	1	1
Snapdeal	1	1	0	0	1	0	1	0	1	0	1	0	0	1	1
Shopclues	1	1	0	0	1	0	1	0	1	1	0	0	0	1	1
Myntra	1	1	0	0	1	0	1	0	1	1	1	0	0	0	0
Koovs	1	1	0	1	1	0	1	0	1	0	0	0	0	0	0
Ajio	1	1	0	0	1	0	1	0	1	1	1	0	1	0	0
TataCliq	1	1	0	0	1	0	1	0	1	1	1	0	0	1	1
UrbanLadder	1	1	0	1	1	1	1	0	1	0	1	0	0	1	1

Pepper Fry	1	1	0	1	1	1	1	0	1	0	1	0	0	1	1
BlueStone	1	1	0	0	1	0	1	0	1	1	1	0	0	0	0
Caratlane	1	1	0	0	1	0	1	0	1	1	1	1	1	0	0

Table 4: Web Design Comparison

	WEB DESIGN							
Websites	Today's Deal/Offer	Previously Viewed Products	Similar Product Suggestions	Frequently Bought Together Suggestions	Season Sale	Mobile App	Trial without Order	Customer Account
Amazon	1	1	1	1	1	1	0	1
Flipkart	1	1	1	1	1	1	0	1
Snapdeal	0	0	0	0	1	1	0	1
Shopclues	0	0	0	0	1	1	0	1
Myntra	1	1	1	1	1	1	0	1
Koovs	0	0	1	0	1	1	0	1
Ajio	0	1	1	0	1	1	0	1
TataCliq	0	0	1	1	1	1	0	1
UrbanLadder	0	0	1	0	1	1	0	1
PepperFry	0	0	1	1	1	1	0	1
BlueStone	0	0	0	0	0	1	1	1
Caratlane	0	0	0	0	0	1	1	1

Appendix – C

Table 5: Paired Similarity Index Matrix

	Amazon	Flipkart	Snapdeal	Shopclues	Myntra	Koovs	Ajio	TataCliq	UrbanLadder	PepperFry	BlueStone
Flipkart	0.780	-									
Snapdeal	0.585	0.561	-								
Shopclues	0.488	0.463	0.439	-							
Myntra	0.683	0.683	0.488	0.415	-						
Koovs	0.366	0.366	0.317	0.293	0.366	-					
Ajio	0.561	0.537	0.415	0.366	0.537	0.366	-				
TataCliq	0.659	0.659	0.537	0.463	0.585	0.366	0.488	-			
UrbanLadder	0.610	0.585	0.512	0.390	0.512	0.366	0.439	0.537	-		
PepperFry	0.634	0.561	0.488	0.366	0.512	0.390	0.463	0.561	0.634	-	
BlueStone	0.439	0.439	0.366	0.293	0.415	0.317	0.415	0.390	0.415	0.415	-
Caratlane	0.537	0.488	0.415	0.366	0.488	0.317	0.439	0.439	0.439	0.415	0.488

Appendix – D Table 6: Ranking of Websites

Websites	Attributes Present	Overall Score	Rank
Amazon	35	8.54	1
Flipkart	32	7.80	2
Snapdeal	23	5.61	6
Shopclues	19	4.63	8
Myntra	28	6.83	3
Koovs	16	3.90	9
Ajio	23	5.61	6
TataCliq	27	6.59	4
UrbanLadder	28	6.83	3
PepperFry	28	6.83	3
BlueStone	20	4.88	7
Caratlane	24	5.85	5

REFERENCES

1. Ankur, P. (2021). E-commerce and Consumer Internet Sector. New Delhi: Ernest and Young.
2. Brohan, M. (1999). Gotta Give it. Internet Retailer (www.internetretailer.com).
3. Erlish, Z. G. (2002). Data Mining by Means of Binary Representation: A Model for Similarity and Clustering. Information Systems FrontiersFrontiers, Vol. 4(2), 187-197.
4. Saikia, H., & Bhattacharjee, D. (2009). Data Mining Into The Websites Of Management Institutes Using Binary Representation. Journal of Applied Quantitative Methods, Vol. 5 No. 1, 53-63.