Study of data science in Online Behavioural Advertising

Ms. Neha Sureshchandra Maurya

Research Scholar
Master's in Information Technology
M. L. Dahanukar College of Commerce
Mumbai University

Abstract - **Since the COVID-19 pandemic spread, social media has become an increasingly important element of people's life. Online activity in the digital domain has expanded considerably, showing a higher reliance on social media platforms for communication in both personal and professional situations.**

Coronavirus has a global impact on e-commerce and hence transformed the nature of business. Despite the Covid-19 issue and economic slowdown, the Indian e-commerce industry exhibited an upward trend after the lockdown, registering a 17% increase in the order volume as of June 2020, when compared to the pre-lockdown period, which continues after the post lockdown period as well. [1]

As a result, advertisers took the time to rethink their methods in order to develop the perfect commercials and reassess who their target audience was. The outbreak proved to be the ideal time for this, given how markets and algorithms are constantly shifting, particularly while individuals are stuck at home.

Companies began looking for ways to attract buyers for their products utilizing data science methodologies in order to enhance their revenue.

We are consolidating our results on how advertisers use data science methodology to acquire customers and the influence on them.

Keywords- **Data science, advertisement, online behavioural advertisement, process model, marketing, behavioural targeting**

I.   INTRODUCTION

Due to the rising usage of electronic devices such as mobile phones and television, online behavioural advertising fuels the internet, allowing businesses to target customers more precisely by studying consumer behaviour on browsing or shopping histories.

Behavioural advertising involves tracking consumers' online behaviours over time in order to provide advertisements tailored to their inferred interests. Behavioural advertisers develop user profiles, categorize user interests, and offer adverts based on demographics and assumptions about user interests using data gathered from internet tracking. Interest categories might be broad or specific depending on the advertiser. Interest categories are used by advertisers to choose and deliver adverts that are relevant to those categories.

When discussing online advertising, several terms are used: demographically targeted, location-based, behavioural, and interest-based advertising. All of these are variations on behavioural advertising.

This could sometimes be beneficial to consumers because it would provide a more personalized online experience, but it could also be manipulative.

In this paper, we provide an overview of targeting approaches and investigate customer perceptions of online behavioural advertising on websites.

II.   Data Science in Advertising and Marketing

Data science is the interdisciplinary study of huge amounts of data using current technology. It aspires to give a thorough, detailed, and sophisticated assessment of raw data. It allows the firm to focus on insights that will have an immediate impact on how the company functions, assists you in making practical predictions for the future, and allows us to make smart marketing decisions. [2]

Marketing, in its most basic form, is the process of discovering client needs and deciding the best way to meet those needs. Advertising, on the other hand, is the practice of promoting a corporation and its products or services through paid channels. Advertising, in other terms, is a component of marketing.

Internet Protocol (IP) addresses; pages visited (on a single website or across sites); length of time spent on pages; advertisements viewed; articles read; purchases made; search terms or other information entered on a site; user preferences such as language and web browser type; and operating system are examples of information that can be

collected in the context of behavioural advertising. Geographical location can also be determined via IP addresses (on the internet) or Global Positioning Systems (GPS) when using mobile communication devices. [7]

The following points outline how data science is applied in marketing: [2]

- Optimization of Channels
- Customer Segmentation
- Lead Targeting and advanced lead scoring
- Real-time interaction and analytics
- Content strategy
- Sentiment analysis
- Maintaining Customer Loyalty
- Predictive analysis
- Recommendation engines
- Marketing Budget Optimization
- Regression Analysis
- Price recommendation for retail

III. Results or Findings

To obtain consumer perceptions of online behavioural advertising, we conducted semi-structured close-ended questionnaires.

A questionnaire is a research instrument that consists of a set of questions or other types of prompts that aims to collect information from a respondent. A research questionnaire is typically composed of both closed-ended and open-ended questions.

Initial questions on knowledge about online behavioural advertising were based on questions on targeting techniques, tracked data, and stakeholder benefit in a study of user opinions and knowledge about online behavioural advertising.

This survey was conducted among 27 employees between the ages of 20- 30 and the results were as follows: more than 51.9% of people spend 2 to 4 hours each day online on average. However, 55.6% of consumers are unaware of behavioural advertising targeting online. The final responders guide is listed in Appendix A.

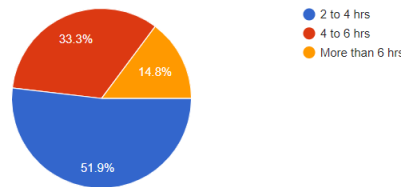How many hours(approximately) do you spend online in a day?
27 responses

- 2 to 4 hrs
- 4 to 6 hrs
- More than 6 hrs

33.3%
14.8%
51.9%

Figure 1

Nonetheless, responders were frequently confronted with a scenario. When they saw advertising for appealing products, they couldn't tell if it was a coincidence or "very clever retargeting."

Did you come across a situation where you have searched a product on some website and later accidently or coincidently you found that through advertisements?
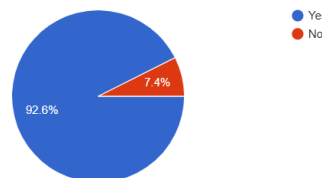27 responses

- Yes
- No

92.6%
7.4%

Figure 2

IV. Methodology

Websites, advertisers, and ad networks are the three key participants in the behavioural advertising model. [7]

Simply defined, websites require funds to function, advertisers want to sell items, and ad networks assist in delivering adverts to a specific population of website users. Furthermore, a few very large advertising organizations can track user behavior across the Internet. Many of these advertising networks are controlled by the same organizations that offer a variety of web-based services and have direct contact with users. This means that certain companies have many points

of entry to its users in various capacities, allowing them to collect information about them. They can then leverage the knowledge acquired to further target users. [7]

Clustering - It uses machine learning algorithms to identify similarities in customer data. As a result, clustering should be applied when parsing through huge datasets. This enables you to possibly target clusters across the whole consumer database, allowing you to personalize at a much higher degree and much faster than people. [6]
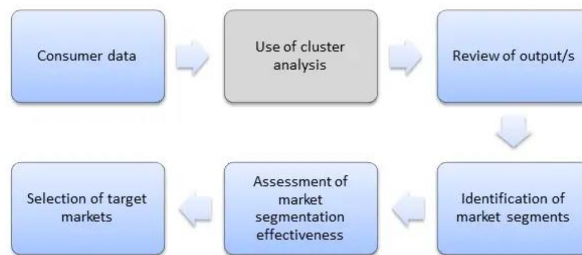

Figure 3: Cluster Analysis [5]

Let's consider an example of an online streaming service that wants to optimize its advertising strategy using clustering:

a.   Data Collection:
Gather data on users, including age, viewing history, preferred genres, and devices used.

b.   Feature Selection:
Choose features like age, genres frequently watched, and the average time spent on the platform.

c.   Normalization:
Ensure that the selected features are on a similar scale to prevent biases in the clustering process.

d.   Clustering Algorithm:
Apply k-means clustering to group users based on similarities in their streaming behavior.
Example: Assume we have three clusters - Cluster A for action movie enthusiasts, Cluster B for users who enjoy comedy and drama, and Cluster C for documentary enthusiasts.

e.   Cluster Profiling:
Analyze each cluster to understand user characteristics.
Cluster A may consist of younger users who enjoy action movies.
Cluster B may include a diverse age group interested in various genres.
Cluster C might be composed of older users with a preference for documentaries.

f.   Ad Targeting:
Design targeted ad campaigns for each cluster.
Cluster A might receive ads for new action-packed releases.
Cluster B might see ads promoting a diverse range of content.
Cluster C could be targeted with ads for upcoming documentaries.

g.   Feedback Loop:
Monitor user engagement and feedback on the ads to assess their effectiveness.

h.   Dynamic Adaptation:
Regularly update the clustering model based on new data to adapt to changes in user behavior and preferences.

By implementing clustering in this manner, the streaming service can tailor its advertising approach to different user segments, increasing the likelihood of users engaging with relevant content and promotions.

Segmentation occurs when a marketer selects criteria to draw specific groups from a big body of data. In other words, it is when you examine your consumer data and select criteria to target a group. [6]

Figure 4: Full STP Process [4]

Let's consider an example of an online fashion retailer using segmentation for its advertising strategy:

a.   Define Segmentation Criteria:

Criteria could include demographics (age, gender), shopping behavior (frequent shoppers, occasional buyers), and style preferences.

b.   Data Collection:

Gather data on users, including age, gender, purchase history, and the types of fashion items they have bought.

c.   Segmentation Analysis:

Use clustering algorithms or statistical methods to identify segments. For instance, users who frequently buy casual wear, users who prefer high-end fashion, and users interested in seasonal trends.

d.   Create Segments:

Group users into segments like "Casual Shoppers," "Luxury Fashion Enthusiasts," and "Trend Followers."

e.   Segment Profiling:

Understand the characteristics of each segment. For instance, the "Luxury Fashion Enthusiasts" segment may consist of young professionals with higher disposable incomes.

f.   Ad Customization:

Tailor advertising content for each segment. Design ads that showcase affordable casual wear for the "Casual Shoppers" and highlight exclusive, high-end collections for the "Luxury Fashion Enthusiasts."

g.   Channel Selection:

Choose appropriate channels based on where each segment is likely to engage. Utilize social media for trend-focused audiences and email campaigns for loyal customers who have made frequent purchases.

h.   Campaign Deployment:

Launch targeted campaigns on selected channels. For example, promote a flash sale on casual wear via social media and send personalized emails to the "Trend Followers" segment about the latest fashion trends.

i.   Monitor and Adjust:

Track the performance of each campaign, analyzing metrics like click-through rates and conversion rates. If the "Luxury Fashion Enthusiasts" segment responds well to a particular campaign, consider extending similar promotions.

By segmenting their audience and tailoring advertising efforts accordingly, the fashion retailer can connect more effectively with different customer groups, enhancing user engagement and driving sales within each targeted segment. The fundamental distinction between the two is that segmentation requires human-defined categories, whereas clustering involves ML-powered groupings. [6]

V.   Advantages and Disadvantages of using Data Science in Online Behavioural Marketing is as follows:
 Advantages: [3]
1.   Broadened worldwide reach.
2.   Personalized campaigns
3.   Lowering costs
4.   Accountability and measurability
5.   Flexibility and adaptability
6.   Increased brand presence and awareness
7.   Easier access to consumer data
8.   Improved customer involvement
9.   Competitor research ability
10.   Ability to execute advertisements around the clock.

Disadvantages: [3]
1.   Ads getting blocked by users.
2.   Increased competition
3.   Inadequate control over ad placement
4.   Ad fraud
5.   Concerns about privacy
6.   There are too many ad possibilities to pick from, and consumers don't know which one to utilize.
7.

VI.   Conclusions:

The research entitled "Study of Data Science in Online Behavioural Advertising" makes an important contribution to the field of Online Advertising and Marketing. This study exhibits data science's ability to attract a large number of customers. The data illustrate how clustering and segmentation can be utilized to significantly enhance sales.

As the area evolves, this research lays the way for additional examination, validation, and implementation of these approaches in a broader advertising and marketing environment.

REFERENCES:
[1] https://www.gapgyan.org/res/articles/(54-57)%20A%20SURGE%20IN%20E%20COMMERCE%20MARKET%20IN%20INDIA%20AFTER%20COVID%2019%20PANDEMIC%20KS.pdf
[2] https://www.netguru.com/blog/data-science-in-marketing
[3] https://adstargets.com/blog/online-advertising-pros-and-cons/
[4] https://www.segmentationstudyguide.com/
[5] https://www.clusteranalysis4marketing.com/a-marketers-guide-to-cluster-analysis/a-marketing-overview-of-cluster-analysis/
[6] https://www.acquia.com/blog/difference-between-segmentation-and-clustering#:~:text=and%20smart%20business.-,What%20is%20clustering%3F,on%20patterns%20in%20their%20behavior
[7] https://www.priv.gc.ca/en/privacy-topics/technology/online-privacy-tracking-cookies/tracking-and-ads/02_05_d_52_ba_01/

Appendix A- Responder Guide
The following questions were asked to the responders.
1. How many hours(approximately) do you spend online in a day?
2. Have you ever heard about Online Behavioural Advertisement Targeting?
3. Did you come across a situation where you have searched a product on some website and later accidently or coincidently you found that through advertisements?
4. What kind of information do websites track?