

Application of Text Mining on Customer Buying Pattern of Cosmetic Products

Athira K, Dhanya M, Sinita S Ashok

Abstract: *The cosmetic industry is one of the major emerging trends in the present world. People are very keen on purchasing cosmetic products, and at the same time, they are very conscious in choosing the right cosmetic items. While making their purchase decisions, customers refer to previous online reviews and the ratings given for that particular product by those who have already purchased, which helps them to get a clear idea about that product. This research paper has focused on looking at the indicators that make the customer's satisfied or dissatisfied towards the purchase-related decisions of cosmetics. Text analysis of the online reviews and the ratings given by customers from the leading e-commerce websites have been analyzed in this paper. The present study focused on understanding some cosmetics attributes that lead to customer satisfaction and dissatisfaction using online reviews. The commonly used attributes from the online text reviews were identified and classified on the polarity basis through the text mining analysis and hence finding out the buying pattern of the customers on cosmetic products*

Keywords: *Sentimental Analysis, Cosmetic Products, Quality, Buying Behavior.*

I. INTRODUCTION

One of the fastest-growing segments of the personal care industry worldwide is the cosmeceutical sector. The global cosmeceutical market is expected to generate \$429.8 billion by 2022, showing a 4.3 per cent CAGR over the 2016-2022 forecast period (Allied Market Research, 2016)[1]. Reports have shown that twelve and six cosmetic products are practised daily by the average woman and man in the United States in pursuit of eternal youth (O'Dell et al., 2016). The changes in people's current habits are having a major impact on the cosmetics industry. Customers are now aware of the use of cosmetics in their everyday lives in an effort to improve their health and overall personality. Cosmetics have a vital role to play in improving the inherent beauty and physical characteristics. Men also use cosmetics gradually in their daily routine, including different types of fragrances and deodorants. This increasing trend in cosmetic products has led to the development of the cosmetics industry worldwide. Personal care product purchase has a high correlation with brand-related attributes. Consumer-brand relationship describes certain brand qualities like brand image, brand attachment, brand love by connection to consumer purchase behaviour. The trend towards organic and natural products

Revised Manuscript Received on April 21, 2020.

* Correspondence Author

Athira K*, Department of Management, Amrita Vishwa Vidhyapeetham, Kochi, India. Email: athira97krishna@gmail.com

Dr Dhanya M, Department of Management, Amrita Vishwa Vidhyapeetham, Kochi, India. Email: dhanyamrajeev@gmail.com

Sinita S Ashok, Department of Management, Amrita Vishwa Vidhyapeetham, Kochi, India. Email: sinitasashok@gmail.com

are rapidly increasing due to safety concerns. This tremendous shift from traditional products to organic products are mostly seen among youngsters. Safety concerns regarding Personal care product are one of the major concern which also affects the purchasing method and consumer behaviour to some extent. A consumer who is well educated follows the ingredients used in any Cosmetic products. A health-conscious consumer relies on recommendations made by doctors, health specialist and skin specialist. Safety concerns occur not only at the point of purchasing the product but also at post-purchase usages.

II. REVIEW OF LITERATURE

The Consumer-brand relationship solely depends mostly on the successful formation of the brand meanings like brand personality, brand attitude, brand association, brand perception and brand image in the minds of customers [2]. Keller's customer-based brand-equity model portrays the consumer-brand resonance, which plays a vital role in any branding. Perception towards the brand can change through experience, influence, or change in brand representation. For this research study, brand image attribute has been taken to study effects on consumer-brand relationship due to safety concerns.

Brand image is the major key driver of brand equity, which refers to the customer's common viewpoint and feeling about a particular brand and has an effect on consumer behaviour [3]. Numerous factors affect the decision-making criteria of a consumer. Brand image is seen as the most compelling factor in consumption decisions [3]. For every marketer, it is essential to keep a brand image to pull than sales to push. Constant public relation activities and creative advertisements are put forth by companies to retain loyal customers.

Regarding the purchase of a product, the theory of regret suggests that the evaluation of the customer post-purchase brand consists of two components: a) element of satisfaction which typically compares the performance of the selected brand with the performance expectations of the pre-purchase and b) an element of regret that is a duty related to the success of the consumer's forewarned brands [4]. This regret experience puts consumer into an emotional state which exhibits lower intention in consumers to repurchase and recommend the brand again. And it leads to higher goals in turning the brand into competing brands and is more likely to engage in brand-damaging behaviours (e.g. complaining, asking for a refund, etc.).

Desai K. (2014) investigated the behaviour of the customer purchasing cosmetic products in Kolhapur. His main aim of the study was to identify the factors influencing the decision of the customer to purchase cosmetic products. The other goal of this research was to study the process of purchasing cosmetic products. The method used to collect data was a convenient method of sampling. Statistical techniques have been used, such as simple percentage and ranking techniques. Quality was the major factor influencing the consumer buying decision, according to his research.

Prof. Anute N., Prof. Khandagle A. Dr. Deshmukh A (2015) studied the behavior of customers purchasing cosmetic products [5]. The main focus of this paper was to research consumer demographic profile and to find out the factors affecting the decision to buy the product. Another important concern was the knowledge of the buying pattern for cosmetic products. They discovered that the majority of people used domestic cosmetic brand, the most successful media for obtaining cosmetic data is television; product value is considered the most important factor in a customer purchase decision.

III. RESEARCH METHODOLOGY

There are primarily two stages in this study. The first step is to gather online reviews and suggestions about various cosmetic products from different websites. The reviews were collected from leading e-commerce platforms such as Amazon, Flipkart, etc., and these platforms were considered ideal as many customers rely on these platforms.

The second step is to use a text mining software called Voyant tool to analyze the data. This software is a web-based program for the study of texts. This technology is used to create word clusters that help us evaluate the frequency of keywords that contain textual data. The keyword available in the tool's context option helps to understand the context in which the most common words are used. The analysis also provides insight into feelings such as positive or negative towards different attributes of service quality. 598 online reviews on various products were collected from e-commerce sites such as Amazon, Flip kart, etc. for this study. The list of products that were considered for review compilation is given in Table- I.

IV. ANALYSIS OF RESULT

The reviews are classified based on positive, neutral and negative reviews, and the frequency table has been obtained as given in Table- II

A. Customer satisfaction or dissatisfaction

To identify customers who are satisfied with the services provided by the product, we did a sentiment analysis on the polarity of the reviews using the python textblob package. Based on the Polarity scores of positive, negative and zero scores, review polarity was identified. The Polarity scores are in such a form that larger scores for positive sentiments. Of the total 598 reviews, 412 reviews were identified as positive sentiment reviews, 151 reviews were identified as negative reviews, and the rest 35 were identified as neutral reviews as graphically represented in Fig. 1. Further, for the analysis of

this study, we only considered the positive sentiment reviews indicating customer satisfaction and negative sentiment reviews indicating customer's dissatisfaction. These reviews were analyzed independently to understand the attributes that most contributed to the overall rating of the products. Hence, the study includes two empirical models, each for customer satisfaction (positive sentiment reviews) and customer dissatisfaction (negative sentiment reviews). Further details of the empirical models are given below in the empirical methodology section.

B. Overall ratings: dependent variable

The dependent variable is the overall rating of the product out of 598 reviews 302 reviews have ratings along with it. Rating of the product is between 1 indicating "terrible" and 5 indicating "excellent" (2 indicates "poor"; 3 indicates "average"; 4 indicates "very good"). Customers who have experienced a positive sentiment would most likely rate the product very good (rating of 4) or excellent (rating of 5) depending on the degree of their satisfaction with the product. Hence, we combined these two ratings together to indicate "higher ratings" for the hotel. Similarly, ratings 3 to 1, indicating average, poor as the overall lower rating.

By analyzing this matrix (Table- III), we could find that out of 185 higher rating reviews (4 and 5), 180 reviews indicate customer satisfaction since their polarity is positive. While 5 reviews have a negative polarity which indicates customer dissatisfaction, this is due to the influence of duplicate products. The words "duplicate", "not satisfactory" made its polarity negative even though they are satisfied with the performance of the product they received some duplicate product which they mentioned in the review made customer dissatisfaction. Out of 117 lower ratings (1to3), 63 reviews have a positive polarity which means the customer is satisfied this maybe because of the fact that rating 3 is also included in this category, so with a neutral opinion is also considered here. A lower rating indicates that the product is not up to the expectation even though it is good. Words like "good", "quality" made its polarity positive, but the rating is low. And 54 reviews have a negative polarity which shows customer dissatisfaction due to bad quality, allergy etc. Therefore their rating is also low. A total of 243 review shows customer satisfaction, while 59 reviews indicate customer dissatisfaction.

A word cluster is formed after the text analysis with Voyants software and is given in Fig. 2.

The context in which the words used are defined and analyzed in order to understand the major factors affecting the customer's purchasing behaviour for all the above words (Table- II).

A linkage diagram containing the most referenced keywords is presented in Fig. 4.

An evaluation was carried out from the context on which keywords are used by the customer (Table- IV).

The study gave an understanding of the customer's buying behavior for cosmetic products which consumers perceive as both positive and negative.

The word "Product" occurred 323 times in the text analysis. Of the 598 reviews received, the word "product" came 323 times, out of which 92 customers had negative product feedback and approximately 217 customers had positive feedback, and the rest were highly neutral.

Some of the context in which this word comes: "the name of the product itself keeps its promise", "amazing color and texture", "very much disappointed with the product as it comes from brand", "never thought that have duplicate", "nice product, good enough to wear", "so its excellence value" etc. It is clear the product is creating mixed opinions among consumers.

In the analysis, a general view emerged that customers need products like lipsticks with amazing colour and texture and it should keep its promises etc.

The next keyword appeared in the analysis was "Skin". The concern about the product that might cause some effects in the skin is expressed here. This word has occurred 277 times. Some of the contexts are given below: "are very cool for Indian skin shades", "kind of irritation or skin damage", "for a bit of dry skin, its good", "good for dry to normal skin, gives good finish for face", "coverage and texture suits sensitive skin", "gives added protection to the skin", "the look become flawless, the skin becomes brighter", etc.

The skin has both positive and negative responses. Of the 598 reviews received, the skin came 277 times and out of this 37 customers had negative skin feedback and approximately 143 customers had positive skin feedback and the rest were highly neutral. The purchasing behaviour of products is also influenced by its effect on the skin like whether it is suitable for every skin type or particularly for specific skin type etc.

The other term that appeared in the review several times was "face". The word face came 123 times. Some of the contexts are: "leaves my face with smooth and glowing", "it blends the face with colour with so well and gives a nice look", "applied it on my face and it looks heartbroken", "chemicals with any other face cream available", "cannot use it as a face mask", "my hair and face was dry for so many days", "it gives a compact rose powder effect", "a lot of potential uses from face to hair masks and cleaning", "I feel it's a good face cream".

There are both positive and negative reactions to the face. Of the 598 reviews collected, 123 customers had negative feedback on the face, and nearly 76 customers had positive feedback on the face, and the rest were highly neutral. Face creams are also of high demand. And also products making a face with smooth and glowing have higher demand in the market.

Another important word that originated from the word cluster is "dry". The word dry occurred 65 times and some of the contexts were: "so nice and does not dry your lips", "in winter apply moisturizer to dry skins", "a blessing for dry skin", "suits very well for dry skin", "good foundation, doesn't dry skin", "quite patchy and dry flakes off quickly", "have oily skin, this will dry out like a boss", "the lotion on let it dry and rub the skin", "good, makes your skin drier, quite disappointed with branded products".

There are both positive and negative dry reactions. Of the 598 reviews collected, 65 came and out of this, 17 customers

had negative feedback on dry, and nearly 46 customers had positive feedback on dry, and the rest were highly neutral. Some products have high demand in winter, which helps customers to get rid of dry lips and dry skin. The extent to which the products are effective in avoiding dry lips and skin determines the demand of the product.

The next important word and one of the most important attributes for assessing purchasing behavior is "Quality". The word "quality" occurred 43 times. Some of the contexts were: "amazing quality with rapid results", "very bad quality", "colours are good but the quality of the product is disappointing", "put more money and get quality product", "is affordable and of good quality", "it is of good quality and its application is very smooth", "cheaper and best in quality than other foundations", "it looked like a low-quality foundation", etc. From here, it is clear that the product is making different opinions that affect customer buying behaviour.

Quality has both positive and negative responses. Out of the 598 reviews, collected quality came 43 times, out of which 20 customers had negative quality feedback and approximately 21 customers had positive quality feedback, and the rest were highly neutral. Customers are really conscious about the quality of the product. Customers expect high-quality products.

The next important word and one of the measuring attributes of buying behaviour are "buy". This word occurred 70 times. Some of the context where this word occurred are: "Whenever I go to buy lipstick I am myself surprise to see even after searching all the brands I would finally choose 9 to 5", "might not use it and buy fresh from store", "please don't buy this products online", "waste of money, please don't buy it", "the one I buy is usually waterproof and thick", "better to buy cosmetic from trusted store than online", "don't buy...it's a duplicate product...!!!", "I won't suggest people buy it", "worth to buy...good product." Etc. From here, it is clear that the customer is having a different kind of opinions. There are both positive and negative responses to buy, as well. Out of the 598 reviews collected buy came 70 times and out of this, 23 of the customers had negative feedback on buy, and 45 customers had positive feedback on a purchase, and the rest were very favourable. Customers buying behaviour depends on many factors like quality of the product, ingredients of product etc.

The next important word and one of the measuring attributes of buying behaviour is "time". This term has taken place 72 times. Some of the context in which this word took place is: "this time I got a duplicate product", "will last for a long time by prevailing the shade", "the eyeliners were delivered on time, but one bottle was broken", "it takes too much time to dry out", "but this is the first time I am not satisfied with this product", "waterproof, smudge-proof but after some time it got faded", "buy it for the first time, it felt good", "colour stays for a longer time".

There are also positive and negative responses to Time. Out of the 598 reviews collected, the time came 72 times and out of this, 13 customers had timely negative feedback, and nearly 57 customers had timely positive feedback, and the rest were highly neutral. Delivery

time of a product, lasting time of cosmetic product all matters in case of buying products.

V. CONCLUSION

Based on the research conducted by text mining of customer reviews on different products, the results showed that buying behaviour of customers is determined by many keywords such as "quality", "face", "dry", "time", and "buy", "product", "skin". From the analysis, we could find that the factor quality is having almost same number of positive and negative reviews, which shows that the factor quality have high influence on the buying pattern while the other keywords like "face", "dry", "time", "buy", "product", "skin" have more positive reviews than the negative reviews. In the case of the keyword "face", the consumers more prefer the products that make a face smooth and glow similarly they don't prefer the product that causes dryness and damage to the face. While during the winter season the customer prefers products that help to avoid the dryness of lips and skin and also the customers are keener towards the products which are suitable for all the skin types and also which does not cause any irritation to the skin.

Consumers are also analyzing the products based on the long-lasting feature of certain cosmetic. Products like lipstick and eyeliner should not smudge and should last for a longer period of time. The on-time delivery of products is also a major concern of customers. They will only buy their products which are worthy for their money spend, that is, which is having good quality assurance and which meets the Customer expectation.

From all these above-listed keywords, we came to the conclusion that there are many factors influencing purchasing behaviour such as affordability, availability, quality etc. From this study, we also could find that in case of cosmetics industry quality is more important for the consumers buying behaviour for certain products amongst the others and also availability factor such as delivery time ease of access is also matters. The customer expectation is met when the quality is good, and most of the customers are convinced by the level of quality assurances given by the product. The presence of organic content in the product makes it more preferable and trustworthy for the customers. Therefore organic products are gaining the attraction of more customers and also is having a high market in the cosmetic industry.

REFERENCES

1. Cao, Q., Duan, W., & Gan, Q. (2011). Exploring determinants of voting for the "helpfulness" of online user reviews: A text mining approach. *Decision Support Systems*, 50(2), 511-521
2. Chang, P. L., & Chieng, M. H. (2006). Building consumer-brand relationship: A cross-cultural experiential view. *Psychology & Marketing*, 23(11), 927-959.
3. Zhang, Y. (2015). The impact of brand image on consumer behavior: A literature review. *Open journal of business and management*, 3(1).
4. Davvetas, V., & Diamantopoulos, A. (2017). "Regretting your brand-self?" The moderating role of consumer-brand identification on consumer responses to purchase regret. *Journal of Business Research*, 80, 218-227.
5. Anute, N. "Consumer Buying Behavior towards Cosmetic Products" Prof. Nilesh Anute, Dr. Anand Deshmukh, Prof. Amol Khandagale.
6. Ravindran, A., & Rejikumar, G. (2017, December). An Investigation on Service Quality Attributes of Indian Hotels by Text Mining. In 2017

APPENDIX

Table- I: List of products

	LIST OF PRODUCTS
1	Lipstick
2	Eyeliner
3	Face Cream
4	Body Lotion
5	Liquid Foundation
6	Rose Powder
7	Eye conic Kajal
8	Eye shadow
9	Mascara
10	Eyebrow Pencil

Table- II: Types of reviews

Positive	412
Neutral	35
Negative	151

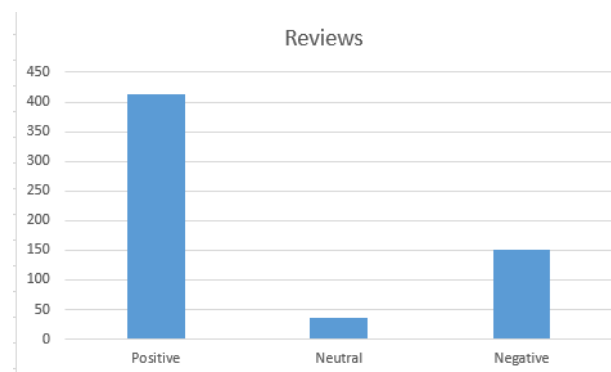


Fig. 1. Reviews

Table- III: Matrix

	Overall higher ratings (4 and 5)	Overall lower ratings (1,2 and 3)	Total number of reviews
Customer satisfaction	180	63	243
Customer dissatisfaction	5	54	59
Total	185	117	302



Fig. 2. Word cluster from text analysis

Table- IV: Words from text mining and its count

WORD	COUNT
Product	323
Skin	277
Good	191
Use	132
Face	123
Time	72
Dry	65
Quality	43

Frequency of Words

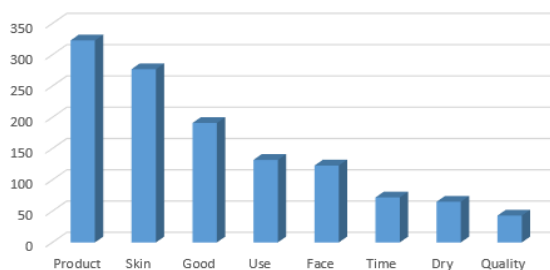


Fig. 3. Frequency of Words



Fig. 4. Links from the text mining

AUTHORS PROFILE



Athira K is currently pursuing MBA at the renowned Amrita Vishwa Vidhyapeetham, Kerala. She wishes to make her area of expertise in Marketing and Analytics. After securing her Bachelor's degree in BSC Mathematics, she turned her attention towards world of Marketing and Analytics to showcase her business and leadership skills. Her leadership traits was seemed and developed during academic carrier by being responsible for coordinating various academic activities. She has actively participated in many well-known business skill



development programs across the state during her academic carrier. Her personal aim is to enter the teaching world and pursue her passion of becoming a lecturer.

Dr. Dhanya M is an Assistant Professor (Senior Grade) in the department of Management, Amrita Vishwa vidyapeetham, Kochi, India. She has graduated her Master of Science (M.Sc.) in Statistics and also holds a Master's degree in Business Administration (MBA). She obtained her UGC-NET in Management and M. Phil. in Statistics. She received her PhD degree in Statistics on the topic "Estimation of reliability measures of some heavy tailed life time distributions" from Mahatma Gandhi University, India. Her areas of interest include Statistical Inference, Bayesian Estimation, Operations Research and Business Analytics. She has published research papers in reputed national and international journals. She is a member of the Kerala Statistical Association.



Sinita S Ashok is currently pursuing MBA at the renowned Amrita Vishwa Vidhyapeetham, Kerala. She wishes to make her area of expertise in Analytics and Operations. After securing her Bachelor's degree in BCA, she turned her attention towards world of Analytics and Operations to showcase her business and leadership skills. She has actively participated in many well-known business skill development program across the state during her academic year. Her passion towards It field had motivated her more to move into analytics world.