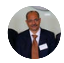
Predicting the consumer's purchase intention of durable goods: An attribute-level analysis

Article *in* Journal of Business Research · December 2017

DOI: 10.1016/j.jbusres.2017.11.031

Sujoy Bag
Indian Institute of Technology Kharagpur



Manoj Kumar Tiwari
Indian Institute of Technology Kharagpur

Predicting the consumer's purchase intention of durable goods: An attribute-level analysis

Sujoy Bag, Manoj Kumar Tiwari*, Felix T.S. Chan

The final version of this manuscript is available at https://www.sciencedirect.com/science/article/abs/pii/S0148296317304770

Highlights

- An attribute level decision support prediction model has been developed.
- Social network mining and sentiment analysis are employed.
- Both linear and nonlinear regression analysis are engaged in the proposed model.
- The outputs of each attribute are compared with four instances.
- Managerial Insights are illustrated and also sensitivity analysis is performed.

Abstract

Recently, Retail 4.0 is progressively demanding the accurate prediction of consumer's purchase intention. In this regard, an attribute level decision support prediction model has been developed for providing an influential e-commerce platform to the customers. In order to build the prediction model, brands' social perception score and reviews' polarity are computed from social network mining and sentiment analysis, respectively. Afterward, an appropriate regression analysis and suitable instances have been identified for each attribute to predict the appropriate product attributes. One of the key findings, the camera attributes: sensor, display, and image stabilization pursue the customer attention at the end of the search. The outcomes of this analysis can be beneficial to e-commerce retailers and prepare an efficient search platform for the customers to obtain the desired durable goods in an adorable form. Finally, the sensitivity analysis has also been performed to test the robustness of the proposed model.

Keywords: online search, consumers review, sentiment analysis, social perception score, linear regression analysis, nonlinear regression analysis.

1. Introduction

Online shopping tendency is meritoriously boosting after the advent of bricks-and-mortar retailers. In the year of 2016, e-retailers have generated the estimated revenue of 1.9 trillion U.S. dollars (7.4% of total retail sales) from 1.61 billion customers globally. Amazon, the leading international e-retail company, has more than 310 million active customer accounts who bought near 136 billion U.S. dollars' goods in 2016 (Statista, 2017). In the first month of demonetization, the growth of digital payment in the world third purchasing power parity country (India) was escalated 271% and simultaneously the cash on delivery was dropped about 30-40% (Chronicle, 2017). Furthermore, out of the total online market, consumers approximately purchase 34% of durable goods (Sen, 2013). Thus, an analysis of online consumer's buying behavior of durable goods is a vital aspect in e-commerce market to represent the online shopping in an eloquent way.

Consumers conduct an extensive search before the purchase of durable goods. As per Bronnenberg et al. (2016), consumer searches camera on an average of 14 times before the purchase. Initially, they search the product based on their needs to gain experience and thorough screening of reviews is exercised by them before a confirmation of purchase. Later, a goal-oriented customer goes for the deeper search to extract the attribute level information and read the associated reviews to make a fruitful decision. Tracking the melody of consumers' online purchasing behavior and storing it in a structural form is a stimulating work. ComScore is a US based leading company that does the job and stores it as comScore Panel Data. However, they have overlooked to store the consumers' screening reviews patterns in their database. In the recent era, an enormous amount of consumers online review data is turning into an interesting and valuable research area for exploring the influential factors in the digital marketing domain. People have given nearly 35 million online reviews up to March 2013 on a single retailer Amazon. It is found that they have increased the revenue of 2.7 billion dollars (Spool, 2009) after setting the question "Was this review helpful to you?" on each customers' review records. Investigating the effect of consumers review data for influencing the purchase intention is a crucial work for providing the prominent online purchasing platform to the customers. Brightlocal.com (BrightLocal, 2016) has found out that 84% of individuals trust the online reviews before purchase. Online reviews impact on consumers' confidence for purchasing a product as well as to provide the real-life experiences, whereas product company benefits from the product feedback to improve the quality of goods.

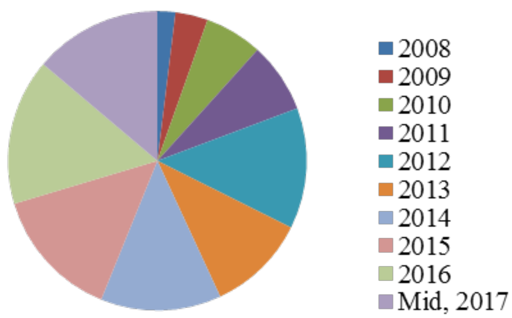
It is essential to explore the consumers' behavior on attribute level for durable goods. Particularly, the question "which search pattern, brand perception, and reviews data are responsible for influencing the consumers' choice?" has become an exclusive query of recent time in online shopping market. In the favor of raised query, different influential factors have been categorized by many researchers in the past for purchasing an online commodity (Chen et al., 2016; Malc et al., 2016). Social network enhances the consumers' perception on brand name (Godey et al., 2016) and also helps the customer to recognize new branded products. Furthermore, brand names impact the consumers' mind for selecting and the willingness to pay for an individual product (Lacroix & Jolibert, 2017; Lim et al., 2016). In this regard, the social perceptual score for the brand's eco-friendly and luxuries nature has been used in the proposed prediction model. Likewise, consumer online reviews influence the customer decision for buying goods (Banerjee et al., 2017) and assists the product company for forecasting the product sales (Fan et al., 2017). An in-depth observation has been carried out by using linear regression analysis for seeing the attribute level consumers' search and choice patterns (Bronnenberg et al., 2016). To address the aforementioned query, researchers investigated the consumers' search patterns, generated the social perceptual score (SPS) (Culotta & Cutler, 2016), and detected the influence of online reviews for purchasing goods. However, the combined effect of attribute level consumers' search and screening reviews patterns on purchasing products has not yet been explored. Further, social perception score is an important aspect in this recent era which should be incorporated in the brand prediction model. Our analysis considers the SPS and encounters its usefulness, and further investigates the joint effect of consumer online search and screening reviews on the purchase decision. According to our best knowledge, another limitation in the existing literature is that no one performed the regression analysis based on the linear and non-linear property of the attributes to deal with consumer buying behavior. However, it is observed that their diverse searching interests on distinct attributes vary with both the linear and nonlinear patterns in the collected data set. This issue has been addressed by simultaneously considering linear and nonlinear regression analysis based on the consumers' searching nature for the specific attributes.

The present research develops an attribute level prediction model to deal with four types of attribute level product selection strategies: (1) search to choose, (2) view reviews to choose, (3) search attributes with looking products' overall reviews to choose, and (4) search attributes with screening corresponding attribute's reviews to choose. Further, predicted attributes values are searched on products database to recommend the relevant products those customers desired to purchase. In the previous study, a number of researchers (Banerjee et al., 2017; Hsu et al., 2017) have considered the influence of consumers' review on purchasing products whereas they have overlooked the reviews which do not affect the customers' intention towards purchasing a product. For example, in this study, insights on the purchase of durable goods have been analyzed by considering digital camera and it is found that the purchase intention for the features Sensor and SLR are independent of customer reviews. This research is important for investigating which attributes and reviews are significant for changing the consumers' mindset towards purchasing any product. Furthermore, the total number of consumers' online search has been normalized to ten equal deciles and influential deciles for each attribute are clearly demonstrated to analyze the insights pertaining to attributes' characteristics.

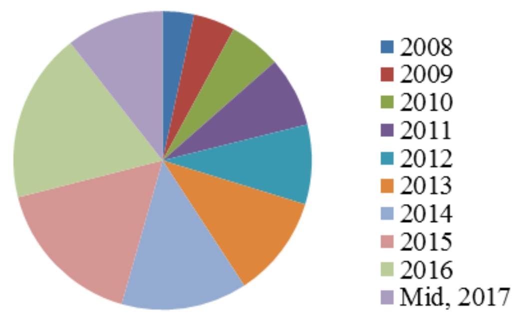
Rest of the paper is organized as follows. In section 2, background and related work are discussed in detail. The case description is shared in section 3. Data collection approach with detailed data description is provided in Section 4. Further, section 5 elucidates the proposed research methodology. The results of prediction graph, influential search deciles, and managerial insights are shown in section 6. Finally, the paper is concluded with limitations and future extensions in section 7.

2. Background and Related Work

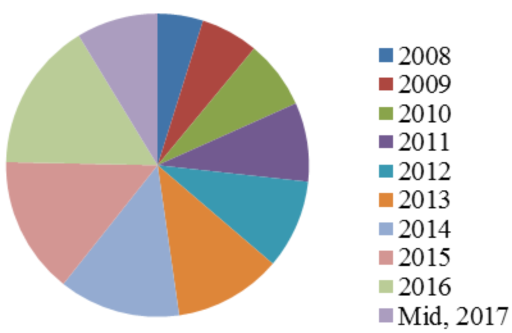
A number of studies have been performed for analyzing the insights of online consumers buying behavior. However, only a few of them have addressed the customers buying behavior for durable goods. Still, an attribute-level prediction model with the integration of consumer search pattern, social perceptual score, and online reviews has not been addressed in the existing literature. By the critical examination of the background and related work, this paper classifies the literature into four subsections, (1) Influential Factors of Online Shopping, (2) Social Network and Brand, (3) Impact of Online Consumers Review, and (4) Predict the Consumers Purchase Intention and Recommendations. Figure 1 shows the research growth of each classification. It seems that the research growth in the related domain is continuously booming markedly. The collection procedure of the records is followed by a simple technique. The keywords of each classification are searched on Scopus database with the constraint 'Article' in document type to collect the number of research articles published in Scopus indexed journals. The aim of this finding is just to show the importance of this research domain.



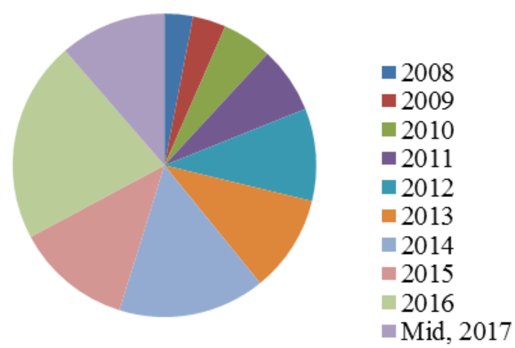
(a) Influential Factors of Online Shopping



(b) Social Network and Brand



(c) Impact of Online Consumers Review



(d) Predict the Consumers Purchase Intention and Recommendations

Figure 1: Published Articles in Scopus Indexed Journals

2.1. Influential Factors of Online Shopping

A number of key factors exist in the literature which influences the consumers toward an online purchase. Primarily, Consumers get influenced for buying a particular branded product from three main online information sources, eWOM (electronic word-of-mouth), manufacturer/retailer, and neutral/third party (Baker et al., 2016; Chen et al., 2016). The gender differences and products type also impact of purchasing digital and non-digital goods (Otterbring et al., 2017; Pascual-Miguel et al., 2015). As per Malc et al. (2016), the price fairness is not only influenced the consumers to buy a product but also spread a negative perception about the seller. Some online resources like video blogs (Lee & Watkins, 2016) changed the consumers' mindset on the physical and social attractiveness of luxury brand perceptions and attitude homophily on para-social interaction (PSI). Lichters et al. (2016) have observed that the compromise effect for buying durable goods is robust than the fast-moving consumer goods (FMCG).

2.2. Social Network and Brand

Social media largely influences the brands quality and loyalty in the consumers' mind (Ibrahim, Wang, & Bourne, 2017). Approximately 80% of the Twitter users are habituated to mention the brand's name in their tweets (Nagy & Midha, 2014). The researchers, Godey et al. (2016) worked on a problem of how social media marketing impact on consumer perception for the brand name of a particular product in a luxury sector. Likewise, Ibrahim et al. (2017) analyzed the social trends on brand image and recognized the different patterns of engagement between customers and company. Brand's individual and social luxury value affect the consumers' willingness to pay for the product (Lim et al., 2016). In favor of measuring the luxury perception from image interactivity, Beuckels & Hudders (2016) were organized an online questionnaire survey of 185 participants. The growth of the luxury good's consumption is exponentially increasing due to the income inequality, social and political instability (Hamelin & Thaichon, 2016), and it varies with different societies, groups, and cultures. Moreover, Lim et al. (2016) identified the paramount factors of consumer buying behavior and compared the luxury vs. regular sportswear brands. In spite of analyzing the effect of luxury brand on purchase intention, Lacroix & Jolibert (2017) have been studied on a luxury watch brands and evaluate the personal legacy value. Gunawan & Huarng (2015) applied fuzzy-set qualitative comparative (fsQCA) analysis and structural equation modeling to examine the social influence and source credibility those are a useful source of information to lead the consumers' purchase intention. To know the brand-consumer relationships in social media, Culotta & Cutler (2016) have developed an automated method to generate the social perceptual score for the perceptual attributes eco-friendliness, nutrition, and luxury. In some instances, consumers' regret after the purchase of a product that affects the consumer-brand relationships (Davvetas & Diamantopoulos, 2017) and their satisfaction influence on product's brand identification.

2.3. Impact of Online Consumers Review

In the recent era, online review systems are making the biases on social influence and product selection. In order to reduce the biases, Askalidis et al. (2017) have differentiated and investigated the retailer promoted reviews and self-motivated reviews for the same product. Further, Hsu et al. (2017) and Maslowska et al. (2017) explored the effects of online consumers review, goods type, and the perceptions in the decision-making process on consumers buying intention. They derived that the negative reviews impact more on purchase decision than the positive reviews, and suggested the retailing management for delivering a quick response to negative comments. Initially, consumers get influenced by the quality of product information (Filieri, 2015). Then, items rating and overall status make the customer decision towards the purchase. According to Fan et al. (2017), previous sales data and consumer reviews are helpful for forecasting the product sales, which was verified by integrating the model of sentiment analysis and Bass/Norton model. Kangale et al. (2016) have been observed the same findings by summarizing the reviews based on the different feature. Consumers perceived risk, usefulness, structural assurance, effectiveness, and so on from the product reviews and get influenced to buy goods individually or in the group (Elwalda et al., 2016; Shi & Liao, 2017; Zhao et al., 2017). Further, Lee et al. (2016) have used consumer reviews to build the mining perceptual map which provides a practical vision for smartphone companies to take suitable marketing decision. In the present study, a number of machine learning based algorithm has been applied on product reviews to classify the reviews into multiple feature vectors (Liu et al., 2017). For example, Banerjee et al. (2017) have proposed a logistic regression-based prediction model to classify the reviews into two sets, high and low trustworthiness. In order to investigate the consumers' screening reviews activities. Luan et al. (2016) introduced an eye-tracking technique and finds that most of the consumers focus on attributes level reviews. Besides, processing and extracting the knowledge from the huge amount of online consumer review data enable scholars to explore the research area of knowledge management systems, big data analytics, and natural language processing (Erevelles et al., 2016; Khan & Vorley, 2017; Pauleen & Wang, 2017). In this research, sentiment analysis, an application of natural language processing, has been performed to transform a huge amount of consumers generated contents to sentimental score and mine the consumer opinion.

2.4. Predict the Consumers Purchase Intention and Recommendations

In the recent digital environment, e-business companies widely used the products recommendation and online advertising for raising the product sales. Google and Rival Microsoft spent nearly 350 and 746 million US dollar respectively in the advertising of its products and services (Sullivan, 2017). A novel active learning approach has been developed by Deodhar et al (2017) to improve the prediction accuracy of recommendation system and search advertising. People usually seek online products with/without any specific target. Ozkara et al. (2016) have presented that e-commerce detects two types of search information such as uncertain and goal-oriented. A study has been conducted on 1261 Dutch car owners by clustering the pre-purchase information search and modeling the structural equation in identifying the sequence of search (Rijnsoever et al., 2012). The search data of 109 participants, those who have purchased at least single water bottled, has been collected (Roscoe et al., 2016) and evaluated the influence in product choice. It is also observed that when participants visit the environmental, economic, or health-related websites where emphasized the benefits for bottled water are more likely to buy. Moreover, Dutta & Das (2017) surveyed the online information search behavior for purchasing mobile phone and laptop on 643 participants from London and Birmingham. They have predicted the influential factor by using multiple linear regression analysis. Similarly, Poel & Buckinx (2005) proposed a prediction model using Logit regression to find the customer purchase intention in their next visit to the retailer website. Initially, they identify 92 input variables and categorized them into four categories such as general clickstream, customer demographics, detailed clickstream, and historical purchase behavior and finally choose 9 variables by using forward and backward variable selection techniques. Jun & Park (2016) categorize the physical product into durable, nondurable and industrial goods and then analyze the correlation between consumers' web search and purchase behavior. The result shows that predicting the purchase of durable goods is significant from the search traffic. However, they exempted the direction of which search path deals with the consumer towards purchase. A number of machine learning based prediction models have been developed with the help of several textual features like polarity, entropy, subjectivity, and reading ease (Ngo-Ye et al., 2017;

Singh et al., 2017) those mitigates the Matthew effect and detects the helpfulness of customer reviews (Mohammadiani et al., 2017; Qazi et al., 2016).

In the past research, the mutual effect of online search and screening reviews data is hardly seen due to the complexity of tracking and linking that information. This research carries out the combined influence of online search and screening reviews for predicting attributes level consumers' purchase intention. The waves of searching and screening consumers' review on attribute level have also been incorporated in this research separately to detect the influential factors in customer purchase decision. In addition, predicting the brand name is difficult from the history of searched brands keyword. Therefore, in this study, the brand's name has been converted into a numerical form of a social perceptual score for the brand's eco-friendly and luxuries nature. This social perceptual score helps the customers to pick up personalized branded product and solves the starvation problem of new branded or less popular products. The aim of this exploration is to assist the consumer to choose their desired products quickly and to help the e-retailing company for taking a suitable managerial decision. An attribute level prediction model has been proposed to achieve the key objectives of this study.

3. Case Description

In this paper, a real-time consumers'online search and choice scenario have been captured for predicting their purchase intention of durable goods. Amazon has been selected as an e-commerce platform for collecting the consumers' real-time search and reviews data. Furthermore, the camera has chosen as durable goods for attribute level analysis. Generally, Amazon predicts the consumers' purchase intention and recommends the products in several ways on their website. Recommendation system having desire to make the application simple and user-friendly in the e-commerce domain and the research on this area is still highly active (Wei et al., 2017). Amazon uses item-based collaborative filtering in their recommendation systems (Smith & Linden, 2017), which has been performing well since 2003. However, existing recommendation systems have few limitations where some of the drawbacks are especially for durable goods. Consumers spend a long period of time in searching the products and screening reviews before purchasing any durable goods. Therefore, capturing the consumers' long-term search and screening reviews data and then storing it in an analytical form for predicting their purchase intention is a challenging task. An attribute-level prediction model has been developed with both linear and non-linear regression analysis for predicting the consumers' purchase intention. Moreover, the sentiment analysis has applied to convert the reviews into polarity and, then incorporate the reviews' polarity with the search patterns. Besides, new items or less popular branded products suffer from cold start problem (Wei et al., 2017) which causes starvation. The proposed model is employed the social network mining to determine the brand's social perception score and to pursue the starvation problem. Proposed case scenario presents in Figure 2 where four types of attribute level product selection strategies are shown.

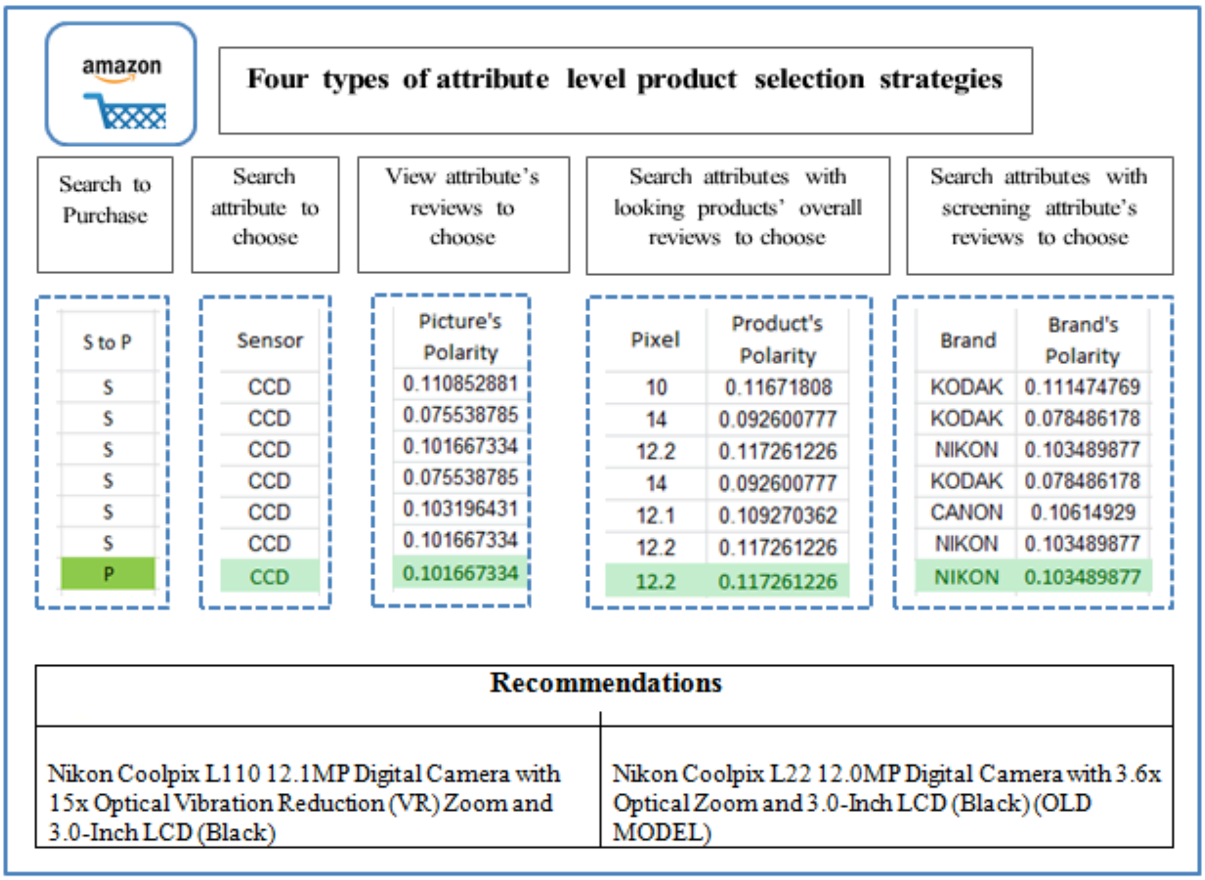


Figure 2: Proposed case scenario

4. Data Collection and Description

4.1. Sources

Initially, three months (October–December, 2010) of total 29069 cameras related consumers' search data (Bronnenberg et al., 2016), and exemplars' follower's IDs (Culotta & Cutler, 2016) have been collected from the informs pubs online repository. Overall 74 and 110 exemplars' accounts have been used to determine the brands SPS (Social Perception Score) for the perceptual attributes of eco-friendly and luxury. Afterward, the irrelevant camera attributes are removed to clean the dataset. Further, Python programming language is used to collect another two types of secondary data. Brand’s follower's IDs and consumers’ opinion are extracted from the brand’s Twitter account and Amazon’s product review pages, respectively. Finally, the reviews data of each product have been cleaned and stored in a different text file.

4.2. Collection Approach

Estimated 70% of customers are extensively searched in a single e-retail website before the purchase (Bronnenberg et al., 2016). Here, Amazon has taken into account for capturing the consumers’ purchasing behavior, which contains whole 3832 search and choice records. The customers have chosen total 769 unique products after an extensive search (i.e. in three months’ time intervals). Each product has Amazon Standard Identification Number (ASIN) which helps to identify the product uniquely. This study fetches total 53 thousand reviews data of listed product IDs. The reviews are collected with a time constraint that all reviews should be on or before December 2010. Beautiful Soup 4, a third party Python library, has applied on Amazon’s products review page to pull out the reviews data. It is observed that the average number of words per review was 165. The records of null product reviews are excluded from the database. Sometimes, people search the product and leave the website without purchasing any product. Consequently, those records are also specifically removed from the database in this study. Finally, total 2472 search records have been selected where 723 unique products are involved. Corresponding twitter account name of each brand are manually collected from twitter website and subsequently, Python programming and Twitter API are used to extract their followers’ IDs.

4.3. Integration and Final Sample Selection

A sentiment analysis has been involved to find two types of reviews’ polarity score, such as product polarity and attributes level polarity. Calculated polarity scores are integrated with the respective search patterns. A social perceptual score of the attributes eco-friendly and luxury are computed for each brand and further replace with the brand's name. All recorded values are normalized into 0-1 scale. Finally, all 18 attributes are separated in 4 instances such as search to choose, view reviews to choose, search attributes and read product’s overall reviews to choose, and search attributes with viewing corresponding attribute’s reviews to choose. It is found that total 301 consumers have bought the camera in the recorded time period. Later, the search path has been normalized into 10 equal deciles. Here, the records of consumers' searches minimum 4 times or in 2 deciles before purchasing any products are only being considered.

5. Research Methodology

This research deals with three components such as attribute level search, product verification from consumers' opinion, and item choices. A framework has been developed to predict the search patterns which are associated with purchase for a particular attribute. The aim of the proposed model is to identify the personalized products and to float the effective advertisement or product recommendation to the customers. In order to transform the online consumers' purchase behavior in a structured form, sentiment analysis, and social network mining are subjected to prepare the training and testing dataset of the prediction model. Consequently, an attribute level prediction model has been developed with the help of linear and nonlinear regression analysis to predict the chosen attributes. The detailed description of the process is organized in subsequent sections:

5.1. Sentiment Analysis

Sentiment analysis is implemented in many cases, Liu et al. (2016) have determined the Twitter sentiment of TV shows in the form of polarity to predict the TV ratings. Vader, a simple rule-based model for sentiment analysis of user-generated contents (Hutto & Gilbert, 2014; NLTK, 2017), has been applied in the present work to classify the text at three different levels, positive, negative and neutral. In the present research, document as well as attribute level analysis is being applied. The polarity of the complete document has been determined by the document level sentiment analysis. Another way, the attribute level sentiment analysis has been split the complete document into the sentences and scanned to identify various relevant interest, and finally have applied sentence-level sentiment analysis. The sentiment score of the corresponding sentence is hence assigned to the corresponding attribute. The selection of the attributes has been carried out based on the frequency distribution of the words from the corpus containing a complete set of camera related reviews. These attributes are further categorized into different aspects based on the similar features.

In present research, it is noticed and presented in Figure 3 that consumers have given higher preference to the camera attributes, Brand, Model, SLR, Sensor, Display, Pixel, Picture, Memory, User-friendliness, and Zoom as compared to the attributes, Movie, USB, Battery, GPS, Face Detection, Image Stabilization, and HDMI in the time of review any product.

In this study, Vader rule-based model (Hutto & Gilbert, 2014) provides the information about what percentage of the document is positive, negative and neutral. The polarity, as well as intensity, is calculated for each attribute by,

$$Polarity = [Positive\_Polarity - Negative\_Polarity] \ldots (1)$$

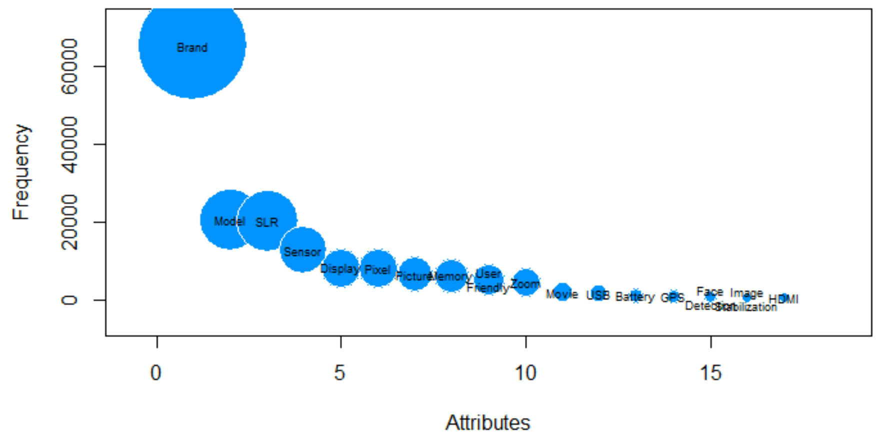


Figure 3: Frequency of different attributes

Here, a basic observation has been presented to give a sense of variability between various searched attribute polarity and chosen attribute polarity. ASP denotes the average polarity of a specific attribute which individual consumer searched several times before deciding to purchase, whereas CP present the polarity of the finally chosen attribute. Figure 4, shows the percentage of times ASP was found to be greater than the CP for each attribute.

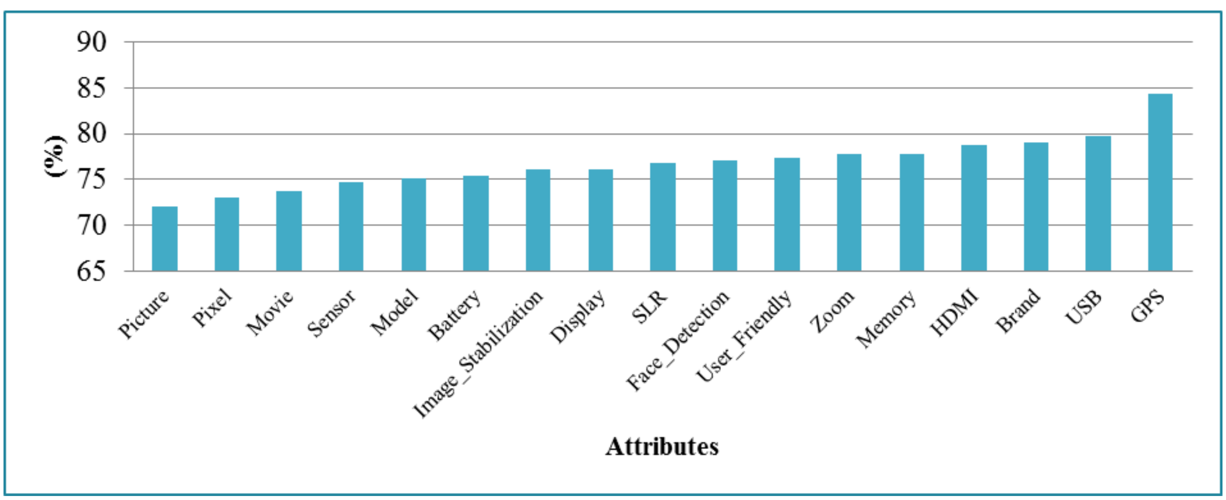


Figure 4: ASP > CP

It has been observed that more than 72% of times the chosen attribute polarity is greater than the searched attribute polarity for any attributes.

5.2. Social Network Mining

Twitter is the most popular social network, used to find customers' perception on a particular brand for different perceptual attributes. In this study, two perceptual attributes, luxury, and eco-friendliness have been selected for determining the brand's SPS. High Social Perception Score (SPS) for different brands in camera reflects the strong relationship between the attributes and brand. Further in equation 2, Jaccard similarity function is used to compute the similarity between the brands and attributes (Culotta & Cutler, 2016).

$$J\left(F_B, F_{E_i}\right) = \frac{\left|F_B \cap F_{E_i}\right|}{\left|F_B \cup F_{E_i}\right|}. \qquad ...(2)$$

Where F_B and F_{Ei} denote the brand's followers and the followers for various exemplars respectively.

The equation 3 is used to calculate the SPS for each brand (Culotta & Cutler, 2016).

$$SPS(B,E) = \frac{\sum_{E_i \in \boldsymbol{E}} \left(1/\left|F_{E_i}\right|\right) * J\left(F_B, F_{E_i}\right)}{\sum_{E_i \in \boldsymbol{E}} \left(1/\left|F_{E_i}\right|\right)}. \qquad ...(3)$$

It's observed that the brand SPS for a perceptual attribute is highly correlated with the consumers' emotions. SPS has been used in the prediction model rather the brand name to predict the consumers' choice accurately for a relatively new and less popular brand. It also helps in predicting the personalized products.

Figure 5 suggests that the SPS for Panasonic is higher than the other brands. In most of the cases, SPS for the perceptual attribute luxury is higher than that of eco-friendly, which indicates that the camera with luxury features is highly demanded than eco-friendly features.

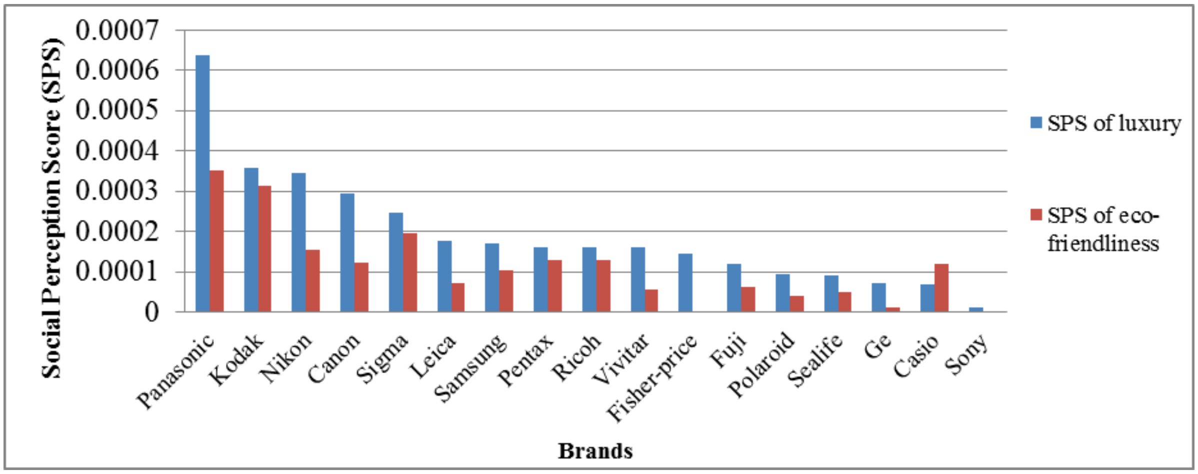


Figure 5: Social Perception Score of luxury and eco-friendliness for different brands

5.3. Normalization for Prediction

Attribute level searches are normalized into ten equal deciles and each decile takes the mean value of the data points in it. Next, the value in all the deciles and chosen attributes are normalized in the scale of 0 to 1.

$$\text{NormVal} = (((OldValue - OldMin) * (NewMax - NewMin)) / (OldMax - OldMin)) + NewMin \quad ...(4)$$

The values in all the ten deciles and chosen attribute serve the purpose of input and output respectively for the regression analysis. Search deciles are selected by the following formula (Bronnenberg et al., 2016):

$$decile(S, N_i) = Ceil\left(\frac{10 * S}{N_i - 1}\right) \quad ...(5)$$

For, S = 1, 2 … N_i – 1, the individual search number, where N_i denotes the total number of the search made by the consumer$_i$.

5.4. Regression Analysis

Regression analysis is employed in this study to predict the attribute level purchase intention of durable goods. Consumers follow the inconsistency for searching and selecting for different attributes, while few attributes consist linearity, and the rest follow nonlinearity. Generally, predicting the chosen attributes which follow the linearity is suitable for linear regression analysis, and similarly, nonlinear regression analysis appropriates for non-linearly followed attributes. In this research, both linear and nonlinear regression analysis has been performed to identify an appropriate model for a specific attribute. In the case of linearity and nonlinearity, multiple regression analysis and GMDH Neural Network are involved, respectively.

5.4.1. Multiple Linear Regressions

Multiple linear regression analysis has been implemented by the statistical software Minitab 17, where search deciles play the role of continuous predictor variables and corresponding chosen attribute as the response variable.

Regression Equation for the Attribute Zoom

$$Output = 0.0133 - 0.0395D_1 - 0.0373D_2 + 0.0897D_3 + 0.1144D_4 + 0.1075D_5 + 0.1017D_6 + 0.0055D_7 - 0.1885D_8 + 0.1151D_9 + 0.7933D_{10} \quad (6)$$

In Table 1 and 2, variance and coefficients of zoom attribute have been analyzed for testing the hypothesis. It is observed from Table 1 and 2 that for zoom attribute deciles 4, 5, 8, 9, and 10 are statistically significant with 90% confidence interval, among them, consumer search on decile 8 negatively correlate with purchase zoom attribute. Similarly, the model adequacy has been tested for the other attributes.

Source	DF	Adjusted SS	Adjusted MS	F-Value	P-Value
Regression	10	5.16126	0.51613	69.06	0
D_1	1	0.00223	0.00223	0.3	0.585
D_2	1	0.00155	0.00155	0.21	0.649
D_3	1	0.01475	0.01475	1.97	0.162
D_4	1	0.023	0.023	3.08	0.081
D_5	1	0.02223	0.02223	2.97	0.086
D_6	1	0.01335	0.01335	1.79	0.183
D_7	1	0.00005	0.00005	0.01	0.932
D_8	1	0.06171	0.06171	8.26	0.005
D_9	1	0.02482	0.02482	3.32	0.07
D_{10}	1	2.09271	2.09271	280.01	0
Error	180	1.34527	0.00747		
Lack-of-Fit	152	1.31194	0.00863	7.25	0
Pure Error	28	0.03333	0.00119		
Total	190	6.50653	*D = Decile		

Table 1: Analysis of Variance

Term	Coefficient	SE Coefficient	T-Value	P-Value	VIF
Constant	0.0133	0.0106	1.25	0.213	
D_1	-0.0395	0.0722	-0.55	0.585	2.51
D_2	-0.0373	0.082	-0.46	0.649	3.68
D_3	0.0897	0.0638	1.4	0.162	2.36
D_4	0.1144	0.0652	1.75	0.081	3.48
D_5	0.1075	0.0623	1.72	0.086	2.29
D_6	0.1017	0.0761	1.34	0.183	2.75
D_7	0.0055	0.0639	0.09	0.932	3.56
D_8	-0.1885	0.0656	-2.87	0.005	2.72
D_9	0.1151	0.0631	1.82	0.070	3.18
D_{10}	0.7933	0.0474	16.73	0.000	1.84

Table 2: Analysis of Coefficients

5.4.2. Non-Linear Regressions Analysis (Neural Network)

The concept of artificial neural network originated from a human neuron that processes information and transmits through electrical and chemical signals. There are three types of layers, an input, output, and a hidden layer. Input layer neurons are passive and take the fixed value of inputs, whereas other layers are active and inputs change with respect to epoch.

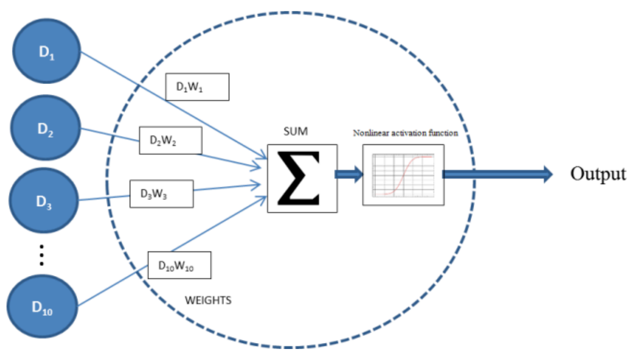


Figure 6: Neural Network Active Node

In Figure 6, search deciles D_1, D_2… D_{10} represent the inputs of the first hidden layer neurons and output to the last hidden layer serves as the input to the output layer and predicts the value of chosen attribute. As shown in Figure 6, the inputs to the neurons in a particular layer multiply with the corresponding weights and sum it up. Further, it has passed through a nonlinear activation function. For instance, $\left[H_1 = f\left(\sum_{i=1}^{10} W_{1i} X_i\right)\right]$, where f is the activation function.

5.4.2.1. Group Method of Data Handling (GMDH) Neural Network

GMDH neural network is a self-organized data mining technique, which can decide the number of variables, structure, and parameters of the model in a self-organizing way. This research has employed the GMDH neural network (GMDH, 2017) for the anticipated problems to predict the search patterns associated with an individual purchase. GMDH neural network used the following Kolmogorov-Gabor polynomial (Ivakhnenko, 1971) model and generates customized polynomial equations for each attribute.

$$\tilde{Y} = a_0 + \sum_{i=1}^{n} a_i D_i + \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} D_i D_j + \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} a_{ijk} D_i D_j D_k + \dots, \qquad (7)$$

Where, $\tilde{Y}$ is predicted output variable, $D = (D_1, D_2, \dots, D_{10})$ is the input vector and a is a vector of coefficients. The Polynomial equation model for pixel attribute has been shown in the appendix section.

6. Results

In this section, prediction graph, influential search deciles, managerial insights and sensitivity analysis are subsequently illustrated with useful findings.

6.1. Prediction Graph

In this study, consumers' purchase intention has been predicted from their search patterns for each attribute. Figure 7 displays the prediction graph of the brand value for social perceptual attribute luxury. The graph shows that the actual chosen attribute data is highly predicted by nonlinear regression analysis.

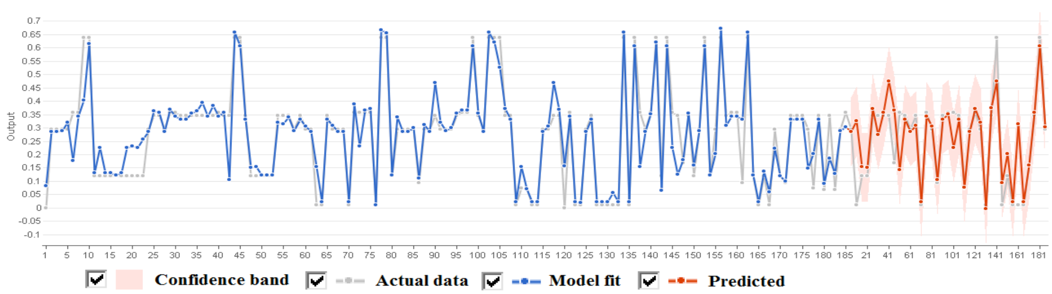


Figure 7: Prediction of the brand value for social perceptual attribute luxury

6.2. Influential Search Deciles

Finding the influential decile for a particular attribute is essential to know the insight of consumers purchasing behavior. Influential deciles have been examined in this study for three instances.

Figure 8 represents the influential deciles of attribute search to purchase a product. It is observed from the Figure 8 that the attributes sensor, display, and image stabilization are effective for purchase at the end of search deciles whereas face detection is prominent in deciles 2, 5, 6, and 10. Moreover, consumers concentrate on searching pixel attribute in the deciles of 4, and 5.

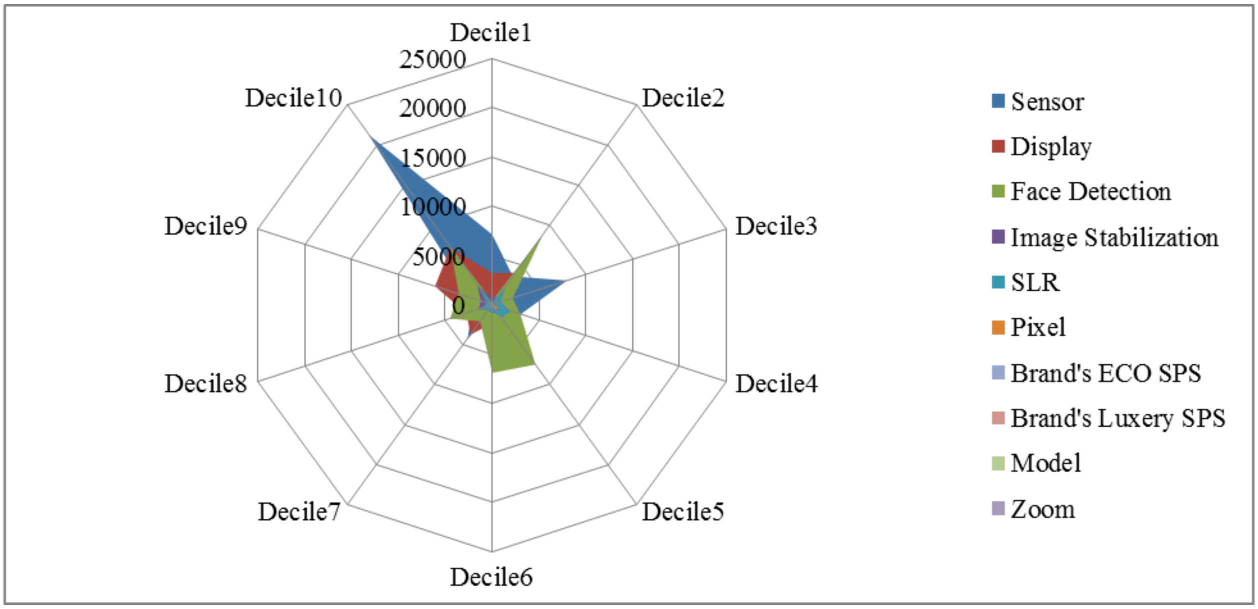


Figure 8: Influential decile of attribute search to purchase a product

The influential deciles for each attribute of screening reviews to buy a product are shown in Figure 9. After analyzed the Figure 9, it is signified that most of the consumers prefer to read the reviews of Memory, HDMI, Movie, Face Detection, Battery, Image Stabilization and picture quality in each of the search deciles.

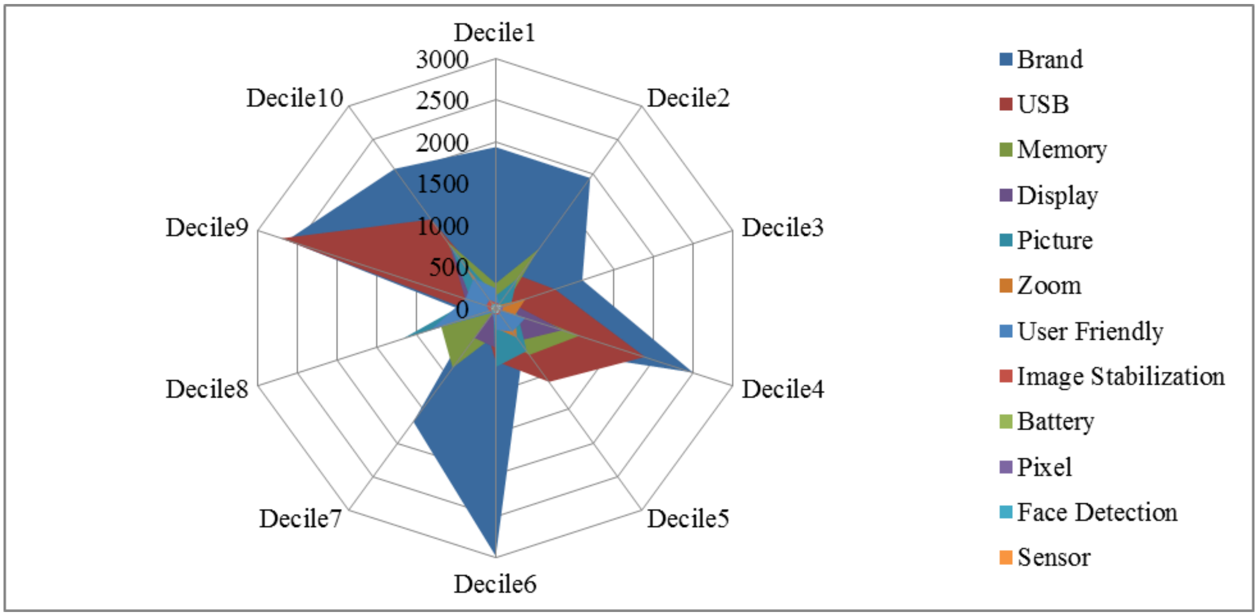


Figure 9: Influential decile of screening reviews to purchase a product

Figure 10 denotes the influential deciles of attribute search with screening reviews to purchase a product. It is noticed that consumers prefer to search and read the corresponding reviews of display attribute in the middle of the search deciles.

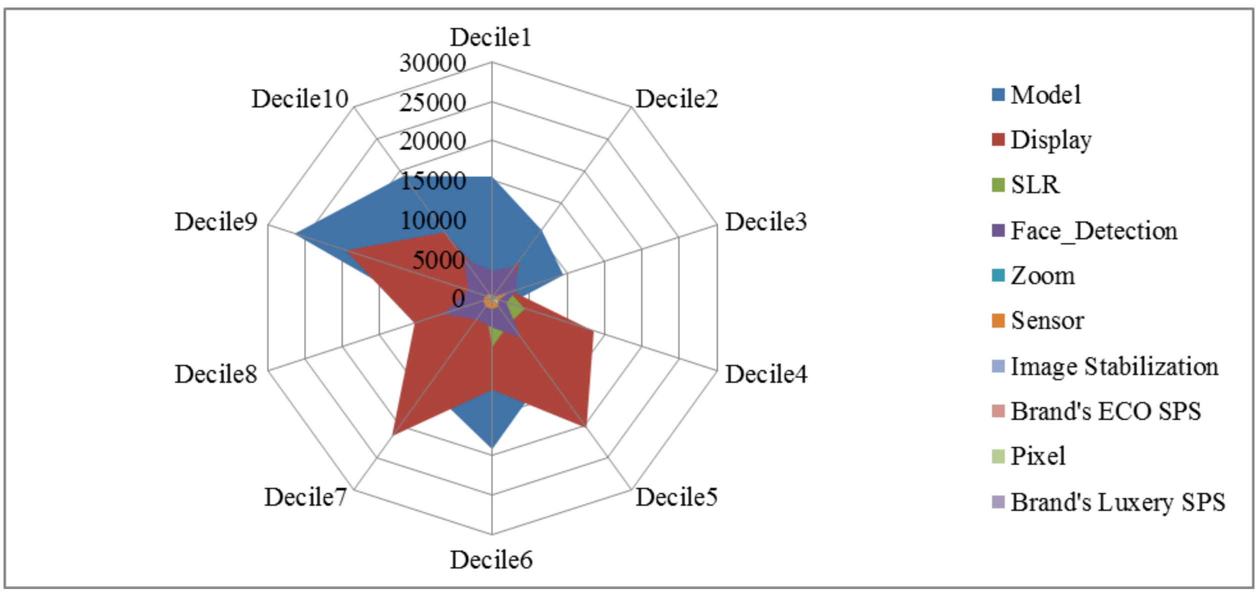


Figure 10: Influential decile of attribute search with screening reviews to purchase a product

6.3. Managerial Insights

Each data set of the attributes has been separated into two parts: (1) 80% for training the model and (2) 20% for testing the efficiency of two regression analyses. Table 3 and 4 show the results of training and testing dataset for the attribute level prediction model. Four instances are categorized in this analysis, where S indicates search attributes towards choose, AR represents attribute level screening reviews to choose corresponding reviews' attribute value, SPR denotes the search attribute with viewing product overall reviews to choose the resultant attribute, and SAR signifies search attribute with screening corresponding attribute's reviews to choose the attribute value. These four instances have been tested with both linear (L) and nonlinear (NL) regression analysis and examined to find appropriate regression analysis for predicting the consumers' attributes choice. Correlation is used to know the relation between attribute's search and choice where -1 represents a strongly negative relation, 0 indicate no relation, and 1 represents a strongly positive relation. The model adequacy has been tested by using the R^2 value where near 1 value represents most significant prediction model. RMSE is the root mean square error which represents the accuracy of the model in percentage. A comprehensive study has been initiated to find the best prediction model for each attribute. Finally, predicted attributes value are searched on products database to recommend the relevant products those consumers intended to purchase.

The variability is obtained in the proposed prediction model for different attributes. The prediction of the attributes brand's SPS for Luxury and ECO (eco-friendly) are highly correlated and strongly significant for nonlinear regression analysis and SAR instance. It indicates that the consumer consciously searches the brand and read the corresponding brand related reviews properly before the purchase. Another interesting observation has been found that prediction of the attributes sensor and SLR are highly correlated with the search and strong significant for the instance S with nonlinear regression, it seems that consumer does not bother about available reviews of the attributes sensor and SLR. Sometimes linear regression performs prominently. For example, the attributes zoom and pixels are relatively strong significant and predict accurately for the instance SPR, a decision can be drawn from here that once a consumer finds appropriate pixel and zoom of a camera and satisfied with product polarity then there are high chances to select the corresponding attribute. Consumers directly check some attribute's reviews before the purchase, among them, memory, picture and user-friendly are relatively strong correlated with the search and significant to predict an appropriate product. The battery is an important feature for purchasing any mobile phone, however, in the case of a camera, the reviews of Battery and HDMI have moderate relation with the search but insignificant for predicting the suitable camera products. Another remarkable observation is that the price attribute has a correlation with the search but not significant for prediction, it suggests to the manager that the consumer can compromise the budget if they satisfied with the searched durable goods. The attribute image stabilization is moderately correlated and significant for all instances. Face detection has a lower correlation with the purchase and weekly significant for predicting the attribute accurately. The attribute movie does not have any correlation with the purchase and insignificant for predict the feature exactly.

Attributes	R	Correlation				R^2				RMSE (%)			
		S	AR	SPR	SAR	S	AR	SPR	SAR	S	AR	SPR	SAR
Brand's SPS of Luxury	L	-	-	-	-	0.83	0.45	0.82	0.81	-	-	-	-
	NL	0.93	0.83	0.93	0.90	0.87	0.70	0.87	0.81	6.46	2.36	3.61	4.08
Brand's SPS of ECO	L	-	-	-	-	0.80	0.45	0.80	0.77	-	-	-	-
	NL	0.92	0.83	0.92	0.88	0.85	0.70	0.84	0.77	4.42	2.36	2.52	2.91
Display	L	-	-	-	-	0.98	0.33	0.90	0.82	-	-	-	-
	NL	0.99	0.67	0.99	0.98	0.99	0.45	0.98	0.97	0.44	2.65	0.90	1.28
Pixel	L					0.67	0.45	0.68	0.67	-	-	-	-
	NL	0.90	0.65	0.90	0.89	0.81	0.42	0.81	0.80	4.81	3.07	2.57	2.84
Price	L	-	-	-	-	0.77	-	-	-	-	-	-	-
	NL	0.98	-	-	-	0.96	-	-	-	1.33	-	-	-
Sensor	L	-	-	-	-	0.74	0.44	0.54	0.48	-	-	-	-
	NL	0.99	0.63	0.82	0.77	0.99	0.40	0.68	0.60	0.03	2.48	0.81	1.03
Zoom	L	-	-	-	-	0.79	0.37	0.79	0.76	-	-	-	-
	NL	0.90	0.71	0.91	0.91	0.82	0.51	0.83	0.83	8.17	2.99	3.86	4.09
Face Detection	L					0.14	0.34	0.13	0.18	-	-	-	-
	NL	0.86	0.53	0.90	0.82	0.73	0.28	0.81	0.67	14.06	4.43	5.87	8.01
Image Stabilization	L	-	-	-	-	0.40	0.37	0.36	0.39	-	-	-	-
	NL	0.70	0.65	0.74	0.65	0.48	0.42	0.55	0.42	35.81	3.91	16.69	19.28
Model	L					0.48	0.27	0.13	0.16	-	-	-	-
	NL	0.66	0.47	0.74	0.73	0.44	0.22	0.54	0.53	2.10	2.86	1.43	1.67
Movie	L	-	-	-	-	0.88	0.17	-	-	-	-	-	-
	NL	0.88	0.42	-	-	0.78	0.18	-	-	11.69	3.92	-	-
SLR	L	-	-	-	-	0.67	0.47	0.62	0.62	-	-	-	-
	NL	0.88	0.70	0.89	0.89	0.78	0.48	0.80	0.79	15.42	2.15	7.56	7.77
Battery	L	-		-	-	-	0.40	-	-	-	-	-	-
	NL	-	0.72	-	-	-	0.52	-	-	-	3.26	-	-
Memory	L	-		-	-	-	0.57	-	-	-		-	-
	NL	-	0.78	-	-	-	0.61	-	-	-	2.27	-	-
Picture	L	-		-	-	-	0.54	-	-	-	-	-	-
	NL	-	0.76	-	-	-	0.58	-	-	-	1.90	-	-
USB	L	-		-	-	-	0.30	-	-	-	-	-	-
	NL	-	0.62	-	-	-	0.38	-	-	-	3.81	-	-
User Friendly	L	-		-	-	-	0.52	-	-	-	-	-	-
	NL	-	0.78	-	-	-	0.61	-	-	-	2.65	-	-
HDMI	L	-		-	-	-	0.25	-	-	-	-	-	-
	NL	-	0.48	-	-	-	0.23	-	-	-	4.01	-	-

Table 3: Results of Training Dataset

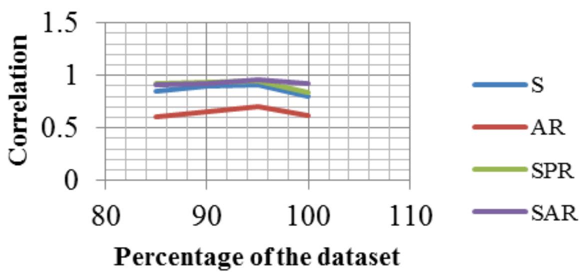
Attributes	R	Correlation				R^2				**RMSE (%)**			
		S	AR	SPR	SAR	S	AR	SPR	SAR	S	AR	SPR	SAR
Brand's SPS of Luxury	L	-	-	-	-	0.81	0.34	0.80	0.78	7.46	**3.24**	4.12	4.16
	NL	0.80	0.62	0.84	**0.92**	0.63	0.28	0.71	**0.85**	9.88	3.35	4.42	3.50
Brand's SPS of ECO	L	-	-	-	-	0.77	0.34	0.77	0.73	5.08	3.24	2.84	**3.01**
	NL	0.79	0.62	0.82	**0.87**	0.60	0.28	0.67	0.76	6.79	3.35	3.38	**3.10**
Display	L	-	-	-	-	0.52	0.18	0.58	0.54	1.82	3.20	1.94	2.76
	NL	0.34	0.45	0.44	**0.68**	-1.79	0.09	0.17	0.46	**1.34**	4.26	1.60	1.81
Pixel	L	-	-	-	-	0.60	0.36	**0.61**	0.60	6.36	3.13	3.39	3.75
	NL	0.60	0.59	0.58	0.58	0.02	0.29	-0.03	0.01	10.10	3.65	5.34	6.07
Price	L	-	-	-	-	0	-	-	-	**3.21**	-	-	-
	NL	**0.53**	-	-	-	**0.21**	-	-	-	5.41	-	-	-
Sensor	L					0.51	0.34	0.47	0.40	0.23	2.45	1.06	1.25
	NL	**0.87**	0.70	0.52	0.77	**0.75**	0.49	0.26	0.57	**0.03**	2.21	1.52	1.22
Zoom	L	-	-	-	-	0.71	0.26	**0.73**	0.70	8.65	**3.59**	4.30	4.91
	NL	0.62	0.32	0.67	0.60	0.28	0.02	0.40	0.27	13.22	4.83	5.97	7.28
Face Detection	L					0	0.24	0	0	23.41	4.69	11.83	11.90
	NL	0	**0.46**	-0.39	0.24	0	**0.18**	-2.76	-2.04	5.37	4.60	3.13	**4.42**
Image Stabilization	L	-	-	-	-	0.34	0.28	0.28	0.33	39.95	4.22	20.72	20.51
	NL	0.62	0.61	0.69	0.62	0.38	0.36	0.47	0.38	38.29	**4.05**	18.23	20.33
Model	L	-	-	-	-	0.16	0.15	0	0	**2.07**	3.00	**2.03**	2.28
	NL	0.21	0.69	0.33	0.22	-0.15	0.43	-0.02	-0.12	2.86	2.99	**2.08**	2.43
Movie	L	-	-	-	-	0	0.09	-	-	21.98	4.10	-	-
	NL	0	-0.07	-	-	0	-0.26	-	-	**3.65**	5.03	-	-
SLR	L	-	-	-	-	0.51	0.36	0.44	0.45	19.32	**2.29**	10.71	10.66
	NL	**0.85**	0.51	0.84	0.80	**0.71**	0.23	0.71	0.64	17.33	2.87	8.66	9.67
Battery	L	-		-	-	-	0.29	-	-	-	3.95	-	-
	NL	-	0.43	-	-	-	0.09	-	-	-	5.35	-	-
Memory	L	-		-	-	-	0.49	-	-	-	**2.51**	-	-
	NL	-	**0.74**	-	-	-	**0.52**	-	-	-	2.74	-	-
Picture	L	-		-	-	-	0.45	-	-	-	**2.12**	-	-
	NL	-	**0.71**	-	-	-	**0.48**	-	-	-	2.44	-	-
USB	L	-		-	-	-	0.19	-	-	-	4.12	-	-
	NL	-	0.65	-	-	-	0.34	-	-	-	3.46	-	-
User Friendly	L	-		-	-	-	**0.42**	-	-	-	**2.91**	-	-
	NL	-	0.52	-	-	-	0.05	-	-	-	3.32	-	-
HDMI	L	-		-	-	-	0.12	-	-	-	4.10	-	-
	NL	-	0.51	-	-	-	0.26	-	-	-	4.00	-	-

Table 4: Results of Testing Dataset

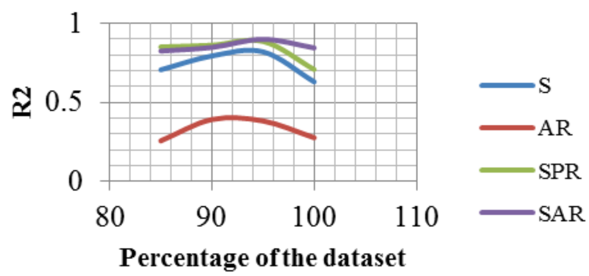
Broadly in this research, the influence of the consumer reviews has been estimated on the product as well as attributes choice. Therefore, managers can easily identify the influential factor of each attribute and can take needful action for the betterment of the e-commerce ecosystem. An efficient e-commerce system can be transferred shoppers into regular customers. Proposed model identifies that reviews maximize the influence of attribute choice and can increase the product demand. Hence, retailers will be benefited in terms of value-added inventory management guided by the demand of the product and attributes.

6.4. Sensitivity Analysis

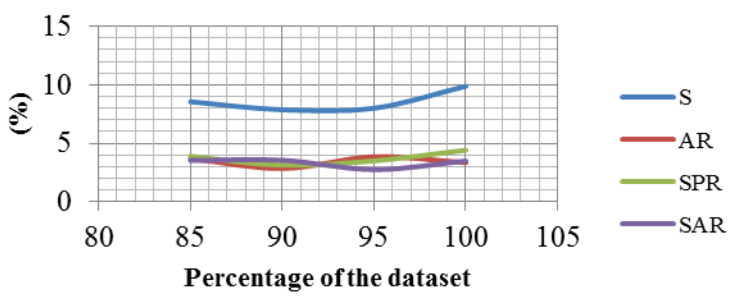
Sensitivity analysis generally performs to test the robustness of any model by observing the variations in result with changing the parameters. In the previous research, many researchers have used the sensitivity analysis in their prediction model. Dag et al. (2017) applied the sensitivity analysis on the prediction model of heart transplantation outcomes through data analytics. Moreover, Maslowska et al. (2017) have engaged the sensitivity analysis to observe the variation of conversion rate while switching the reviews' volume. In this study, the size of the dataset has been varied to test the sensitivity analysis. The results of Brand SPS Luxury Attribute are shown in Figure 11. The instance SAR is performing consistently for any size of the dataset.



(a) Correlation



(b) R2



(c) RMSE

Figure 11: Sensitivity Analysis of Brand SPS Luxury Attribute

7. Conclusions

The goal of this research is multi-directional. (1) Product and e-retailing company could take an abundant managerial decision from this analysis. (2) The prediction model can help the customers for making an effective purchase decision in a shorter time period, and (3) Furthermore, e-retailer could sponsor a personalized product recommendation platform in a lower cost. Finding the consumers' preferred product is a stimulating work. In this regard, an attribute level prediction model has been proposed to identify the customers' choice for each attribute. To obtain the desired products, both linear and nonlinear regression analysis has been performed and compared the output with four instances. Finally, the suitable regression model and instances have been selected for each attribute, and afterward, the combination of predicted attributes values are searched on products database to recommend the relevant products to the customer. Moreover, the sensitivity analysis has also been performed to test the consistency of the proposed model. One of the key findings of the results is the attribute sensor predicted accurately only from the consumer search history. Further, attributes demand can also be easily computed from the proposed prediction model and on the basis of these results, the products inventory can be managed. Furthermore, Amazon or other e-retailing company can apply the proposed model, especially for durable goods in their application to provide a significant product recommendation to the customers.

7.1. Limitations and Future Extensions

Nowadays, the consumers' internet usage strategy has been upgraded; some degree of changeability might be seen from the previous patterns which should think over in the model. In the future, proposed attribute level prediction model can be extended in multiple dimensions. (1) First of all in this study a single category of durable goods is considered, further proposed model can be tested on multiple durable goods those consumers searched in various website. However, the proposed model is expected to work for any categories of durable goods generally. (2) A generic Vader rule-based sentiment dictionary has been used in this analysis to find the reviews polarity. The sentiment analysis can be extended by proposing a personalized sentiment dictionary for each category products. Further, utilitarianism performs a similar role as that of sentiment analysis which can be explored in the future research. (3) Basically, in this prediction model, a pilot study has been evolved by only considering the three months' consumers searched and screening reviews data. In future, big data analysis can be involved for data collection, sentiment analysis, social network mining, and regression analysis separately or jointly to do the analysis in more sophisticated way. Specifically, Apache Spark, an open-source cluster-computing framework, can be performed for extracting, cleaning, storing and integrating the huge amount of user generated contents, and implementing the prediction model followed by sentiment analysis, social network mining, and regression analysis. (4) In addition, some more influential factors like sales, discount, offers, deals, seasonality, etc. can also be incorporated to perform the model more efficiently and predict adequately.

Acknowledgments

The work has been supported by the project of E-Business Center of Excellence and funded by Ministry of Human Resource and Development (MHRD), Government of India under the scheme of Center for Training and Research in Frontier Areas of Science and Technology (FAST), Grant No. F.No.5-5/2014-TS.VII.

Appendix:

Polynomial equation model for pixel attribute

Y1 = -0.00512837 + N65*0.615767 + N128*0.392909

N128 = -0.127143 + N234*0.839793 + N251*0.375303

N251 = -0.3184 + N262*0.892216 + N307*0.646443

N307 = 0.578968 + Decile7*1.02321 - "Decile7, cubert"*0.688378

N234 = -1.01195 + "Decile10, cubert"*1.65898 + N284*0.362115

N65 = -0.119594 + N181*0.876579 + N256*0.325746

N256 = -0.494099 + N284*0.93975 + N319*0.89615

N319 = 0.588588 + Decile9*0.951224 - "Decile9, cubert"*0.653417

N284 = 0.585124 + Decile4*1.08194 - "Decile4, cubert"*0.761281

N181 = -0.115033 + N211*0.877137 + N262*0.317472

N262 = 0.583878 + Decile5*1.10609 - "Decile5, cubert"*0.782693

N211 = 0.0167397 + Decile10*0.968224

References

Askalidis, G., Kim, S. J., & Malthouse, E. C. (2017). Understanding and overcoming biases in online review systems. *Decision Support Systems*, *97*, 23–30.

Baker, A. M., Donthu, N., & Kumar, V. (2016). Investigating how word-of-mouth conversations about brands influence purchase and retransmission intentions. *Journal of Marketing Research*, *53*(2), 225–239.

Banerjee, S., Bhattacharyya, S., & Bose, I. (2017). Whose online reviews to trust? Understanding reviewer trustworthiness and its impact on business. *Decision Support Systems*, *96*, 17–26.

Beuckels, E., & Hudders, L. (2016). An experimental study to investigate the impact of image interactivity on the perception of luxury in an online shopping context. *Journal of Retailing and Consumer Services*, *33*, 135–142.

BrightLocal. (2016). Local consumer review survey. Retrieved from https://www.brightlocal.com

Bronnenberg, B. J., Kim, J. B., & Mela, C. F. (2016). Zooming in on choice: How do consumers search for cameras online? *Marketing Science*, *35*(5), 693–712.

Chen, J., Teng, L., Yu, Y., & Yu, X. (2016). The effect of online information sources on purchase intentions between consumers with high and low susceptibility to informational influence. *Journal of Business Research*, *69*(2), 467–475.

Chronicle, D. (2017). The impact of Demonetisation on E-Commerce. Business, Economy. Retrieved from http://www.deccanchronicle.com/business/economy/030317/impact-of-demonetisation-on-e-commerce.html

Culotta, A., & Cutler, J. (2016). Mining brand perceptions from twitter social networks. *Marketing Science*, *35*(3), 343–362.

Dag, A., Oztekin, A., Yucel, A., Bulur, S., & Megahed, F. M. (2017). Predicting heart transplantation outcomes through data analytics. *Decision Support Systems*, *94*, 42–52.

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*.

Deodhar, M., Ghosh, J., Saar-Tsechansky, M., & Keshari, V. (2017). Active Learning with Multiple Localized Regression Models. *INFORMS Journal on Computing*, *29*(3), 503–522.

Dutta, C. B., & Das, D. K. (2017). What drives consumers' online information search behavior? Evidence from England. *Journal of Retailing and Consumer Services*, *35*, 36–45.

Elwalda, A., Lü, K., & Ali, M. (2016). Perceived derived attributes of online customer reviews. *Computers in Human Behavior*, *56*, 306–319.

Erevelles, S., Fukawa, N., & Swayne, L. (2016). Big Data consumer analytics and the transformation of marketing. *Journal of Business Research*, *69*(2), 897–904.

Fan, Z.-P., Che, Y.-J., & Chen, Z.-Y. (2017). Product sales forecasting using online reviews and historical sales data: A method combining the Bass model and sentiment analysis. *Journal of Business Research*, *74*, 90–100.

Filieri, R. (2015). What makes online reviews helpful? A diagnosticity-adoption framework to explain informational and normative influences in e-WOM. *Journal of Business Research*, *68*(6), 1261–1270.

GMDH. (2017). Expand your predictive analytics. GMDH Shell. Retrieved from https://gmdhsoftware.com/

Godey, B., Manthiou, A., Pederzoli, D., Rokka, J., Aiello, G., Donvito, R., & Singh, R. (2016). Social media marketing efforts of luxury brands: Influence on brand equity and consumer behavior. *Journal of Business Research*, *69*(12), 5833–5841.

Gunawan, D. D., & Huarng, K.-H. (2015). Viral effects of social network and media on consumers' purchase intention. *Journal of Business Research*, *68*(11), 2237–2241.

Hamelin, N., & Thaichon, P. (2016). Consumer motives and impact of western media on the Moroccan luxury buyer. *Journal of Retailing and Consumer Services*, *32*, 164–170.

Hsu, C.-L., Yu, L.-C., & Chang, K.-C. (2017). Exploring the effects of online customer reviews, regulatory focus, and product type on purchase intention: Perceived justice as a moderator. *Computers in Human Behavior*, *69*, 335–346.

Hutto, C. J., & Gilbert, E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Eighth international AAAI conference on weblogs and social media*.

Ibrahim, N. F., Wang, X., & Bourne, H. (2017). Exploring the effect of user engagement in online brand communities: Evidence from Twitter. *Computers in Human Behavior*, *72*, 321–338.

Ivakhnenko, A. G. (1971). Polynomial theory of complex systems. *IEEE Transactions on Systems, Man, and Cybernetics*, *1*(4), 364–378.

Jun, S.-P., & Park, D.-H. (2016). Consumer information search behavior and purchasing decisions: Empirical evidence from Korea. *Technological Forecasting and Social Change*, *107*, 97–111.

Kangale, A., Kumar, S. K., Naeem, M. A., Williams, M., & Tiwari, M. K. (2016). Mining consumer reviews to generate ratings of different product attributes while producing feature-based review-summary. *International Journal of Systems Science*, *47*(13), 3272–3286.

Khan, Z., & Vorley, T. (2017). Big data text analytics: an enabler of knowledge management. *Journal of Knowledge Management*, *21*(1), 18–34.

Lacroix, C., & Jolibert, A. (2017). Mediational role of perceived personal legacy value between consumer agentic generativity and attitudes/buying intentions toward luxury brands. *Journal of Business Research*.

Lee, A. J. T., Yang, F.-C., Chen, C.-H., Wang, C.-S., & Sun, C.-Y. (2016). Mining perceptual maps from consumer reviews. *Decision Support Systems*, *82*, 12–25.

Lee, J. E., & Watkins, B. (2016). YouTube vloggers' influence on consumer luxury brand perceptions and intentions. *Journal of Business Research*, *69*(12), 5753–5760.

Lichters, M., Müller, H., Sarstedt, M., & Vogt, B. (2016). How durable are compromise effects? *Journal of Business Research*, *69*(10), 4056–4064.

Lim, C. H., Kim, K., & Cheong, Y. (2016). Factors affecting sportswear buying behavior: A comparative analysis of luxury sportswear. *Journal of Business Research*, *69*(12), 5793–5800.

Liu, M., Pan, W., Liu, M., Chen, Y., Peng, X., & Ming, Z. (2017). Mixed similarity learning for recommendation with implicit feedback. *Knowledge-Based Systems*, *119*, 178–185.

Liu, X., Singh, P. V., & Srinivasan, K. (2016). A structured analysis of unstructured big data by leveraging cloud computing. *Marketing Science*, *35*(3), 363–388.

Luan, J., Yao, Z., Zhao, F., & Liu, H. (2016). Search product and experience product online reviews: An eye-tracking study on consumers' review search behavior. *Computers in Human Behavior*, *65*, 420–430.

Malc, D., Mumel, D., & Pisnik, A. (2016). Exploring price fairness perceptions and their influence on consumer behavior. *Journal of Business Research*, *69*(9), 3693–3697.

Maslowska, E., Malthouse, E. C., & Viswanathan, V. (2017). Do customer reviews drive purchase decisions? The moderating roles of review exposure and price. *Decision Support Systems*, *98*, 1–9.

Mohammadiani, R. P., Mohammadi, S., & Malik, Z. (2017). Understanding the relationship strengths in users' activities, review helpfulness and influence. *Computers in Human Behavior*, *75*, 117–129.

Nagy, J., & Midha, A. (2014). The value of earned audiences: how social interactions amplify TV impact. *Journal of Advertising Research*, *54*(4), 448–453.

Ngo-Ye, T. L., Sinha, A. P., & Sen, A. (2017). Predicting the helpfulness of online reviews using a scripts-enriched text regression model. *Expert Systems with Applications*, *71*, 98–110.

NLTK. (2017). Sentiment Analysis. Retrieved from http://www.nltk.org/howto/sentiment.html

Otterbring, T., Ringler, C., Sirianni, N. J., & Gustafsson, A. (2017). The Abercrombie & Fitch

Effect: The Impact of Physical Dominance on Male Customers' Status-Signaling Consumption. *Journal of Marketing Research*.

Ozkara, B. Y., Ozmen, M., & Kim, J. W. (2016). Exploring the relationship between information satisfaction and flow in the context of consumers' online search. *Computers in Human Behavior*, *63*, 844–859.

Pascual-Miguel, F. J., Agudo-Peregrina, Á. F., & Chaparro-Peláez, J. (2015). Influences of gender and product type on online purchasing. *Journal of Business Research*, *68*(7), 1550–1556.

Pauleen, D. J., & Wang, W. Y. C. (2017). Does big data mean big knowledge? KM perspectives on big data and analytics. *Journal of Knowledge Management*, *21*(1), 1–6.

Qazi, A., Syed, K. B. S., Raj, R. G., Cambria, E., Tahir, M., & Alghazzawi, D. (2016). A concept-level approach to the analysis of online review helpfulness. *Computers in Human Behavior*, *58*, 75–81.

Roscoe, R. D., Grebitus, C., O'Brian, J., Johnson, A. C., & Kula, I. (2016). Online information search and decision making: Effects of web search stance. *Computers in Human Behavior*, *56*, 103–118.

Sen, S. (2013). Consumer durables account for 34% of Indian online market: Google report. Business Today. Retrieved from http://www.businesstoday.in/technology/internet/

Shi, X., & Liao, Z. (2017). Online consumer review and group-buying participation: The mediating effects of consumer beliefs. *Telematics and Informatics*, *34*(5), 605–617.

Singh, J. P., Irani, S., Rana, N. P., Dwivedi, Y. K., Saumya, S., & Roy, P. K. (2017). Predicting the "helpfulness" of online consumer reviews. *Journal of Business Research*, *70*, 346–355.

Smith, B., & Linden, G. (2017). Two Decades of Recommender Systems at Amazon. com. *IEEE Internet Computing*, *21*(3), 12–18.

Spool, J. (2009). The magic behind Amazon's 2.7 billion dollar question. Retrieved from http://www.uie.com/articles/magicbehindamazon/2009

Statista. (2017). Statistics and facts about global e-commerce. Retrieved from https://www.statista.com/topics/871/online-shopping/

Sullivan, L. (2017). Google, Facebook, Microsoft Spent Hundreds Of Millions To Advertise In 2016. Retrieved from https://www.mediapost.com/publications/article/302460/bing-google-facebook-spent-hundreds-of-millions.html

Van Den Poel, D., & Buckinx, W. (2005). Predicting online-purchasing behaviour. *European Journal of Operational Research*, *166*(2), 557–575.

Van Rijnsoever, F. J., Castaldi, C., & Dijst, M. J. (2012). In what sequence are information sources consulted by involved consumers? The case of automobile pre-purchase search. *Journal of Retailing and Consumer Services*, *19*(3), 343–352.

Wei, J., He, J., Chen, K., Zhou, Y., & Tang, Z. (2017). Collaborative filtering and deep learning based recommendation system for cold start items. *Expert Systems with Applications*, *69*, 29–39.

Zhao, K., Stylianou, A. C., & Zheng, Y. (2017). Sources and impacts of social influence from online anonymous user reviews. *Information & Management*.