


Online Product Reviews and Their Impact on Third Party Sellers Using Natural Language Processing

Akash Phaniteja Nellutla, Symbiosis Centre For Management and HRD, Symbiosis International University (Deemed), India

Manoj Hudnurkar, Symbiosis Centre For Management and HRD, Symbiosis International University (Deemed), India

Suhas Suresh Ambekar, Symbiosis Centre For Management and HRD, Symbiosis International University (Deemed), India

Abhay D. Lidbe, Alabama Transportation Institute, University of Alabama, USA

 <https://orcid.org/0000-0001-7791-129X>

ABSTRACT

The purpose of this paper is to gain insights from the online product reviews of e-commerce sites such as Flipkart and Amazon and analyze its impact on third party sellers. To judge the authenticity of a product, reviews are more useful than ratings, since ratings do not give a complete picture. It is always preferred to consider both the product and seller reviews to have a seamless delivery and defect less product. In this paper, natural processing methods are used to gain insights by considering online reviews of a product. Methods such as sentiment analysis, bag of words model help to understand the impact of online product reviews on the seller's ratings and their performance over some time. The reviews are categorized into positive, negative, and neutral using sentiment analysis. Further, topic modeling is done to find out the topic reviews are majorly referring to. The seller reviews for a specific product after analysis are compared with the overall seller reviews to judge the authenticity. The results of this paper would be beneficial to both the consumers and sellers.

KEYWORDS

Bag of Words, Natural Language Processing, Online Product Reviews, Sentiment Analysis, Topic Modeling

INTRODUCTION

Online shopping has completely changed the way of shopping. In India alone, e-commerce growth is expected to be \$120 billion, growing at an annual rate of 51 percent, by 2020 (IBEF Report, 2020). Online shoppers stand at 25 million in India. These days, it needs just one click and we are just one day away from the product arriving at the doorstep. With technology changing day by day and more options coming in, people prefer shopping online than visiting stores (Xiong et al., 2020). Online giant retailers like Amazon, Flipkart, and Myntra are making it easier by providing various products in various categories throughout the year. Sometimes, the products that are not available in the offline

DOI: 10.4018/IJBIR.20210101.0a2

This article, published as an Open Access article on December 18, 2020 in the gold Open Access journal, International Journal of Business Intelligence Research (converted to gold Open Access on January 1, 2021), is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

stores can be bought online. Also, the price tends to be on the lower side when compared to the store price. These are some of the reasons why people now prefer to purchase online.

Online retailers have a lot of advantages when compared to traditional brick-and-mortar stores. Online retailers can sell as many products as they want. They have no constraints with respect to inventory or space management. These giant retailers have also gained the trust of the consumers by providing them with one-day delivery options, lower prices, trusted sellers, product reviews, seller reviews, and ratings. So a normal customer visits the website, chooses the product, looks at the price and ratings, reads the reviews, and then checks out (Vineet et al., 2020). The most important factors that he/she usually considers are the ratings and the reviews of the product. These two factors play a crucial role while a consumer selects a product. But online shopping has its own problems. There is a chance of users receiving a faulty or another product instead of the one they ordered. Also, the ratings sometimes do not give a complete picture of the value of the product. For example, a consumer purchasing a mobile phone solely based on ratings might have a problem with the battery life of the mobile. The reviews help to sort out such problems. Online reviews are a great source of information and play a crucial role when a consumer purchases a product. The consumer decides whether to go ahead or back down based on the reviews (Wang and Chen, 2020). Reviews are written by the consumers after purchasing a product and using it for a short/long duration. These reviews are written for both the product and the seller. The product reviews majorly define the different aspects of the product and the experience of the customer after using it. The seller reviews talk about the delivery, service provided, and sometimes the behavior of the seller (Xiaolin et al., 2019).

Majority of the past studies talk about online product reviews, their importance, sentiment analysis being used to analyze the reviews, and the impact of these reviews on sales (Ali et al., 2020). These studies have considered customer reviews at a high level or were limited to only product reviews. An analysis has been done on those reviews using different techniques (sentiment analysis, ML algorithms) in order to provide product recommendations and find out the opinion of the customer regarding the product. But the seller reviews are completely ignored. This study considers both the product and seller reviews for further analysis in order to simplify the existing process.

In light of the above aspects, this paper aims to study and analyze the online reviews (both of the product and the seller) using Natural Language Processing (NLP) techniques. NLP was used in this study because it helps in understanding the data, even if it is unstructured, and then takes decisions based on the insights gathered. Also, sentiment analysis, which is one of the most important fields of NLP, was used in this study. Sentiment analysis not only helps in understanding the current and historic context of text but also guides in predicting the future (Shobana and Murali, 2020). Additionally, the information stored in the text can be used as an indicator such as positive, negative, and neutral (Al-Sharuee et al., 2020). This aspect of sentiment analysis is very helpful in this study, which can be used as a signal for decisionmakers. Relating the same to this study, NLP helps to understand the impact the reviews create on the seller and also compare them with the seller reviews in general to judge their authenticity. Finally, the consumers can decide on purchasing the product from a particular seller based on the above analysis and also the seller can improve their functioning based on the customer feedback (Caitlin et al., 2019).

The further sections of this paper such as the Literature Review section discuss the various studies where similar analysis has been done and the gap this study is going to fill. The other sections, namely Research Methodology and Analysis and Results, discuss the methodology used in order to obtain the insights, the analysis done, and the results. The discussion section discusses the results in comparison with other studies in this field. Finally, certain recommendations and limitations have also been included.

LITERATURE REVIEW

This study is mainly related to mining online product reviews, seeing their impact on third-party sellers, and understanding how NLP techniques can be helpful in decision making of the buyer. There are many research papers related to these topics on the internet. Several papers presented their findings, which formed the base for this study.

The online reviews' importance depends on when they are presented (Camilleri, 2020). Online communication on products and services, also known as Electronic Word of mouth (eWOM), forms an important aspect from a marketing point of view. eWOM, which now takes the form of reviews, is important for every firm and has an impact from an economic perspective (Doi and Hayakawa, 2020). The consumer and the business are both affected by online reviews. The reviews can be positive or negative. They can also be conflicting (Kordrostami and Rahmani, 2020). Sometimes the consumer can also give a reaction in the form of a tentative review (Ifie, 2020). When the customer is given the option to express his opinion on the product after purchasing it on an online platform, it is the responsibility of the seller/merchant to respond to his concerns in order to avoid loss of reputation (Obeidat et al., 2017). While positive reviews/5-star ratings help in improving the business of sellers and retailers, negative reviews can cause the same impact in the opposite way. It is therefore important to differentiate between genuine and fake reviews (Wang and Chen, 2020). The businesses are mostly using structured data in the form of customer ratings on a scale of 1 to 5, but it may not provide the real information about whether the customer is saying what he really feels (Gallagher et al., 2019). This is where the unstructured data in terms of customer reviews will be useful.

Latent Similarity Analysis (LSA), TF-IDF model, Support Vector Machine, and Logistic Regression can be used to detect fake reviews (Viji et al., 2020). Purchase verification can also be of some sort of help in this regard (Jiaxiu et al., 2020). Though there is a lot of evidence available that the online ratings have a significant impact on pricing strategies, it is not completely clear. So, from a product customization perspective, firms should take a standpoint instead of waiting for the product reviews (Li et al., 2020). Due to the increased usage of technology and mainly the Internet, people actively seek out other opinions while shopping online. This is where reviews come into the picture. Since their importance has increased, product review analysis has become highly beneficial to predict current trends. Opinion mining and sentiment analysis form a wide area of research (Shobana and Murali, 2020). Sentiment analysis can be explained as *"the process of extracting and classifying opinions expressed in digital documents or web content, and then analyzing the resulting data against well-developed lexicon, a list of words or phrases classified into positive, negative, and neutral"* (Chang et al., 2019, p. 265). Lately, sentiment analysis through NLP is getting much attention in literature. Sentiment analysis is considered as the most appropriate method for textual analysis (Sann and Lai, 2020). Unstructured textual data in the form of comments, reviews, and chats in social media provides an added advantage to understand the business and help business strategies and can be analyzed effectively using sentiment analysis (Ramaswamy and DeClerck, 2018).

The reviews, tweets, and posts on the Internet are usually not structured. It is all unstructured data. But due to advances in the discipline of machine learning, now it is not just about interpreting the text but also about analyzing it and taking decisions (Xianhao et al., 2020). Also, it is not possible to go through hundreds or thousands of reviews given by consumers. This is where NLP can be used. Currently, a lot of research is going on in this area, and it is successful due to an increase in computational power and easy access to data. This field has brought in a lot of meaningful results in areas such as retail, healthcare, finance, media, and human resources. This study also employs the NLP techniques in order to provide the appropriate results.

NLP is an approach that helps to analyze and present textual data for the purpose of processing this information for a variety of tasks (Ali et al., 2020). Sentiment analysis and text mining are the processes for such analysis, which can categorize the statements in the reviews. After the reviews are classified as positive, negative, and neutral, based on review polarity, an unsupervised review

selection is used to understand the produced sentiment pattern (AL-Sharuee et al., 2020). The core ingredient for sentiment classification in online product reviews are adjectives, verbs, and adverbs. In literature, it has been reported that the general adverbs strongly help in classifying sentiments with better accuracy (Chauhan et al., 2020). Nowadays, Deep Learning Modified Neural Networks (DLMNN) and Adaptive Neuro-Fuzzy Inference Systems are being used for sentiment analysis and future prediction of product. These algorithms tend to increase accuracy even more (Sasikala and Sheela, 2020).

These online reviews not only create a significant impact on the customer's point of view but also cause an effect on product sales (Xiaolin et al., 2019). These in turn also cause an effect on the third-party sellers on the retailing platforms (Wen et al., 2020). In light of the gaps in research and the aims of the study, this study explores the answers to the following research questions.

Research Question 1: Are purchase decisions of customers influenced by online product reviews? What are the factors that influence a purchase decision?

Research Question 2: Do the third-party supplier reviews add value to purchase decisions in addition to product reviews?

All the other research studies in this area are limited/confined to presenting the impact of online product reviews by various techniques like gaming rules and other machine learning techniques.

So this research paper mainly focuses on bridging the gap by analyzing the seller reviews and reviews mentioning the seller, apart from the product reviews, using NLP techniques to finally present the insights, understand the impact it creates on third-party sellers, and choose the best performing seller with reference to India (only considering Indian product reviews).

RESEARCH METHODOLOGY AND DESCRIPTION

This study was done in two parts. First, primary research was conducted to understand customer opinions about use of online reviews in purchase decisions. Additionally, the secondary research of collecting and analyzing customer reviews was conducted. A positivist approach of evaluating numerical ratings cannot provide the details of reviews, like reasons for the customers' likes or dislikes. It has been observed that a rating of 4 out of 5 may be very good but the review comments may include complaints about the product or service. Such complex reviews get neglected in a positivist approach. The study assumes that the customer is a complex individual who is unaffected by external promotional efforts and needs in-depth analysis for better understanding of their behavior. Customers tend to speak their mind if provided with an opportunity of a detailed review rather than just a numerical rating (Gallagher et al., 2019). Hence this approach helps in understanding why customers feel the way they feel.

The data collection section describes the procedure followed to gather data. The data extraction section describes the method of scraping the reviews from the e-commerce website. The topic description section of the methodology gives a brief description of all the topics that were discussed and used in this study. Finally, the data modeling section describes the end to end process of building a model.

Data Collection

In the primary research, a questionnaire was floated and nearly 150 responses were collected. The questions were related to the users' preferences while purchasing a product online and their points of view about product and seller reviews. The respondents belonged to various age groups ranging from 20 to 60 years. Each one of them had previous experience of purchasing/shopping online. Many

of them were able to differentiate between product reviews and seller reviews. The results from this research formed the fundamental aspect of secondary research.

In secondary research, online giant retailer Amazon.in was chosen. A product with a high number of ratings and reviews (Sennheiser CX 180) was considered. Reviews from India were taken for analysis. Only one product and that too in particular the above product was chosen because of the availability of considerable number of reviews not only for the product but also for the sellers. In the other products, there was an imbalance in the number of reviews. The reviews of only those users who had a Verified Purchase tag were considered. Not only the tag but also the history of the user, such as number of purchases, his usage, and number of reviews given to date were some other factors taken into consideration.

Data Extraction

Mining in general or data mining is all about analyzing data and gaining insights from the given data. In this study, mining of online reviews was done in order to understand the reference.

The entire process of analysis was carried out using Python and Excel.

Libraries/tools used in Python include: Pandas used for data manipulation and Analysis, NLTK (Natural Language Toolkit), a tool that has a number of libraries related to statistical and symbolic NLP, Seaborn for attractive and informative visualizations, Gensim for unsupervised topic modeling, and pyLDAvis, which is an interactive LDA visualization package.

Web scraping technique was used to scrape the reviews, and a sample of 5,000 product reviews was considered. Web scraping is an automated method that is used to extract the data from websites and save it in a spreadsheet on the local computer (Lusiana et al., 2019). The extracted output in this study consisted of different attributes, such as order number, URL, author, title, date, content, rating, and hyperlink.

This was the data that was used for sentiment analysis.

Topic Description

Natural Language Processing

Natural Language Processing or NLP is related to computer science, linguistics, and artificial intelligence. It deals with the interactions between computers and humans using the natural language. In particular, analyzing large amounts of natural language data by machines to gain insights is what NLP is all about. NLP gives machines the ability to read, understand, and derive meaning from human languages (Vera et al., 2020).

Some use cases include chatbots, identifying fake news, stopping spam, prediction of diseases from past health records, and sentiment analysis.

Also, NLP and text analysis were used in analyzing UK annual report narratives. Though the work was limited only to the English language, further research was done to implement in other languages (Mahmoud et al., 2019).

Sentiment Analysis

Sentiment analysis is a subfield of NLP that is used to identify the emotions, attitude, and evaluations based on computational treatment of subjectivity in a text (Sameh and Ozgur, 2020). Sentiment analysis can provide a lot of information and insights about the choices of the customers and their decision drivers (Shobana and Murali, 2020).

In this study, VADER (Valence Aware Dictionary and sEntiment Reasoner) sentiment analysis is used. This is a lexicon and a rule-based sentiment analysis tool that is used to classify a text as positive, negative, or neutral (Anton and Martin, 2020). VADER analysis also gives a compound score, which is a sum of all lexicon ratings normalized between -1 to +1. In order to classify a text as a positive, negative, or neutral one, Table 1 can be used as a reference.

Table 1. Compound Score Metric

Sentiment	Compound Score
Positive	≥ 0.05
Neutral	$> -0.05 \ \& \ < 0.05$
Negative	≤ -0.05

The VADER sentiment analysis has the following advantages over other traditional methods of sentiment analysis.

It works very well on the social media data, which may be related to many domains; it is fast; it does not suffer from speed performance tradeoff; and it basically does not require any input other than the training data.

Once the reviews are classified as positive, negative, or neutral, corresponding percentages are found out in the entire dataset (5,000 reviews).

Data Modeling

The next step is to extract important and useful topics in each category review. The concept of Latent Dirichlet Allocation (LDA) was used to extract the same (Fan et al., 2016). LDA was used in this study because it provides better accuracy than other unsupervised algorithms such as K-means.

The following steps are involved in the process.

Data Preprocessing

This is an important step in which the punctuations and stopwords are removed and the reviews are normalized as much as possible. The most frequent words in the data are also extracted. Since the frequent words consist of many articles that are actually not relevant, tokenization and lemmatization is done. This helps in extracting the relevant and useful frequent words from the reviews (Guang et al., 2017).

Stop Words

Stop words are the most commonly used words such as “a,” “an,” “the,” “in,” which are usually removed while processing natural language data.

There is no universally followed list for stop words. The list can be built from scratch. Generally, while processing the text using NLP techniques, the common words, which add very little or no value, are removed. Sometimes while removing the stop words, the relevant information useful for our analysis may also be removed. In such cases, a very minimal amount of stop words list should be prepared, depending on the objective the user wants to achieve (Jatinderkumar and Rajnish, 2016).

Tokenization

Tokenization is the process wherein the text/string is converted into a list of tokens/words. This is useful in extracting the relevant words from the reviews.

In the process of tokenization, the text is segmented into words or sentences depending on the command of the user. These divided pieces are called tokens. In the process of segmentation of the text, many characters such as punctuations are removed. It is complicated to deal with texts wherein there are a lot of parentheses, hyphens, and other punctuation marks.

Lemmatization

Lemmatization works by identifying the parts of speech of a given word and then applying more complex rules to transform the word into its true root, known as lemma.

In general, the objective of lemmatization is to reduce the word into its base form and then combine/group together the different forms of the same word. For example, swimming, swims, and swam are all the forms of the word “swim,” so “swim” is their lemma. Lemmatization, depending on the context, can also differentiate between similar words that have different meanings. This way it can even solve problems like disambiguation (Khaltar and Atsushi, 2009).

There is another process known as stemming, which is also used by many. Stemming is also similar to lemmatization but it is less effective than the latter. The approach followed by both techniques is also different. The process of stemming is fast and requires less computational power as compared to lemmatization. But it is not the optimal method. Lemmatization, on the other hand, is a resource-intensive task and requires higher computational power.

In this study, lemmatization was done on the tokenized reviews, which gives the most relevant frequent terms in the reviews and also serves to further remove the noise.

Topic Modeling

In order to discover the abstract topics that occur in a collection of documents, a statistical modeling method called topic modeling was used.

Topic modeling, as the name suggests, is used to identify topics present in a text and also to understand the hidden meaning from that text corpus (Ike and Sathish, 2020).

Topic modeling is used for various other purposes such as clustering of documents, feature selection, unstructured data information retrieval, and categorizing textual data into large chunks.

It also has several other real-world applications such as dimensionality reduction, recommendation engine, and text summarization.

Building an LDA Model

Latent Dirichlet Allocation or LDA is an example of a topic model that is used to classify text in a document to a particular topic. It is one of the most commonly used models. LDA is an unsupervised learning method that builds topic per document model and words per topic model, modeled as Dirichlet distributions (Qing et al., 2020).

To be more precise, LDA builds words per topic model by assigning the words to random topics. The topics are not defined by the user earlier. Only the number of topics he/she wishes to know is given as an input. Then the algorithm will map all the relevant words to a particular topic. Many iterations are done by the algorithm in mapping the words to different topics by taking different cases into consideration. The probabilities that a word belongs to a particular topic and the document will be generated by a topic are calculated multiple times until the convergence of the algorithm.

Once the LDA model is built, it is used to print out the topics it has learned. In this study, the LDA model was built using the document term matrix as one of the parameters. Bag-of-words can also be used as a parameter instead of the same.

Bag-of-Words

Bag-of-words model is the most commonly used model that gives the count of words in a piece of text. The count of each word is given in the form of an occurrence matrix. It does not consider the grammar and order of words (Hong et al., 2016).

Then the topics are visualized in a 2-D space. The visualization is interactive and it displays the topics it has learned, along with the most relevant words in each topic.

This process was followed for all three categories—positive, negative, and neutral reviews, in order to find out the topics in which the product excels and the ones in which it lags. This would help the consumer while deciding on the product.

In the next step, the reviews consisting of the seller keyword, more particularly the name of the seller are scraped in the same way. The same process as above is followed in order to categorize into positive, negative, and neutral. Then topic modeling is done in order to identify the areas where the

Table 2. Age distribution of respondents

Age	Percentage
Below 18	2.70%
18 - 30	66.90%
31 - 50	21.40%
50+	9%

service/delivery was good or went wrong. Then these reviews and also the positive, negative, and neutral percentages of the specific product chosen are compared with the particular seller reviews that are given by consumers for various products in general. This validates whether the seller is authentic in their delivery or not, depending on the recent and overall performance. This comparison of seller reviews on a granular level helps the consumers to choose the seller for the specific product depending on the above factors. If the product reviews mentioning the seller match with the overall seller reviews on a performance basis, then the consumer can go forward with the seller. In the other case, they can choose a seller with better ratings and better performance match percentage. This enables the consumer to not choose the seller solely based on ratings and price (lower price than other sellers). This also helps the seller to analyze what went wrong with the delivery and service of the product and also improve their performance if lagging.

DATA ANALYSIS AND RESULTS

Of the 145 responses received from the primary research, 71 were from females and 74 were from males. Of them, 97 belonged to the age group 18–30, 31 to the age group 31–50, 13 to the age group of 50+, and 4 members were below 18. This is shown in Table 2.

Around 58 percent of respondents shopped online occasionally and 35 percent shopped often. Only 1 person never shopped online. This helps to understand the rise of e-commerce. This is shown in Table 3.

Though there is a small chunk of people who prefer only online shopping, it can be stated that people tend to prefer selecting the product virtually than going to the store and trying it.

In the case of online product reviews, around 91 percent marked that they read the reviews before checkout. They tended to read around 12–15 reviews on an average while purchasing a product online. This can be observed in Table 4.

Out of the above 91 percent, nearly 75 percent also went through different seller options available and their reviews. The majority of people consider product reviews to be more important than the seller reviews, as seen in Table 5. Though it seems to be true considered from a consumer point of view, the seller ratings and reviews should also be given equal importance.

Table 3. Frequency of shopping online

Shopping Frequency	Percentage
Always	5.90%
Often	35.20%
Occasionally	57.90%
Never	1%

Table 4. Respondents views about Product Reviews

Read Reviews	Percentage
Yes	91%
No	3%
Maybe	6%

Table 5. Importance of different attributes according to respondents

Category	Low Importance	High Importance
Price	74	71
Product Reviews	5	140
Product Ratings	15	130
Seller Ratings	52	93
Seller Reviews	50	95

In Table 5, it is clearly visible that the seller ratings and reviews are also given low importance, and it is greater when compared to product reviews and ratings.

Considering all the above-stated facts, secondary research was done by scraping the product reviews for a particular product—Sennheiser CX 180Street II in-ear headphone (black), without mic—from Amazon India. The selected product had around 58,000 ratings and 48,000 reviews. A sample consisting of 5,000 reviews was used for analysis.

VADER sentiment analysis was used to categorize the reviews as positive, negative, and neutral. The compound sentiment score metric in Table 6 was taken into consideration and then the reviews were classified.

Out of 5,000 reviews, 81 percent (4,054) were positive, 15 percent (736) were negative and 210 (4 percent) were neutral.

The positive reviews were considered for further analysis. The reviews were then tokenized and then lemmatization was done in order to get the frequent words used in positive reviews.

Table 7 shows the top 30 frequent words before the stop words are removed.

The top 30 frequent words after the stop words are removed are shown in Table 8.

Tokenization and lemmatization are done in order to get more relevant frequent words in the extracted reviews.

Some of the frequent words include “quality,” “awesome,” “noise cancellation,” “warranty,” “clarity,” “treble,” and “volume,” as shown in Table 9.

LDA model was built in order to get the top 5 relevant topics based on the above positive reviews. This topic model can be visualized in a 2-D space and each topic also has its own relevant words. The words shown in Figure 1 are more inclined toward the functioning topic of the headphone.

Table 6. Compound Score Metric

Sentiment	Compound Score
Positive	≥ 0.05
Neutral	$> -0.05 \ \& \ < 0.05$
Negative	≤ -0.05

Table 7. Top 30 frequent words in Positive reviews

Word	Count		Word	Count
is	5500		with	1900
the	5200		but	1800
and	4000		are	1800
for	3200		bass	1750
I	3000		not	1700
this	2900		have	1650
it	2900		Earphones	1500
a	2850		i	1450
to	2800		that	1450
of	2740		very	1450
in	2300		product	1400
quality	2200		best	1400
sound	2100		The	1400
you	1900		my	1000
good	1900		price	900

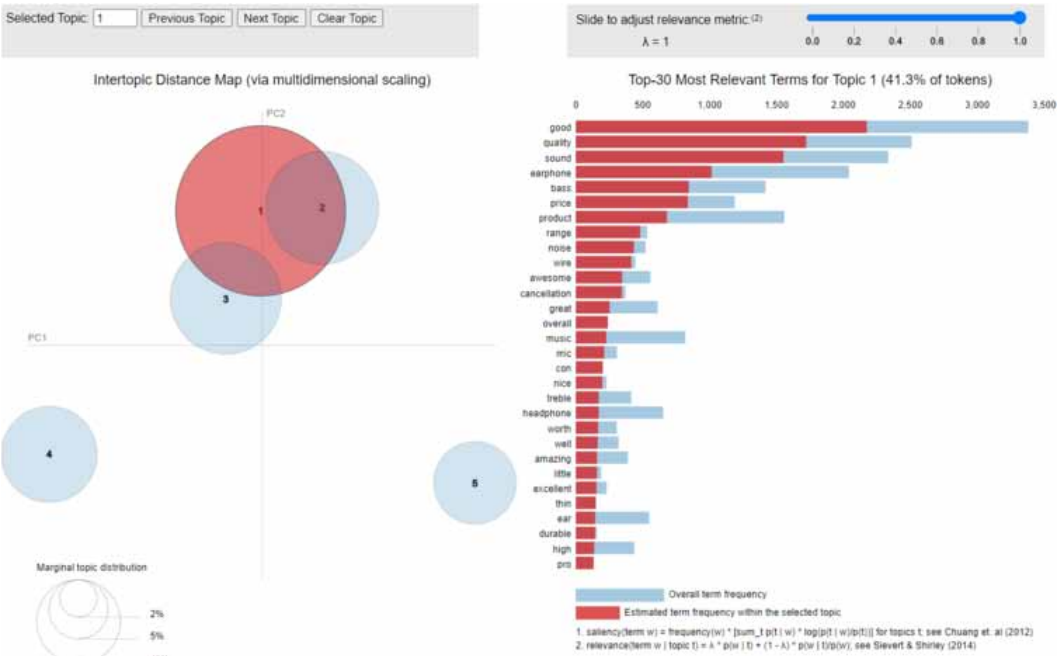
Table 8. Top 30 frequent words in positive reviews after stop words are removed

Word	Count		Word	Count
sound	2700		awesome	600
quality	2600		mic	550
good	2500		using	540
bass	2200		like	540
product	1550		range	500
best	1500		better	500
earphones	1450		buy	450
price	1250		but	450
one	1100		also	450
the	1050		headphone	450
earphone	1000		noise	450
music	800		use	450
Sennheiser	750		years	430
great	700		really	430
ear	600		amazing	410

Table 9. Top 30 frequent words in positive reviews after Lemmatization

Word	Count		Word	Count
good	3500		month	750
quality	2700		range	745
sound	2400		noise	740
earphone	1900		high	650
product	1600		time	640
bass	1500		amazing	630
price	1300		phone	450
year	900		cancellation	440
music	850		well	430
ear	830		warranty	430
great	820		clear	410
sennheiser	810		clarity	400
headphones	810		treble	400
awesome	800		mic	300
wire	760		volume	290

Figure 1.



This process was followed for the negative reviews also. The categorized negative reviews were considered and the relevant frequent words were extracted using tokenization and lemmatization.

LDA model was built in order to get the top 5 relevant topics based on the negative reviews. The topic model can be visualized in a 2-D space and each topic also has its own relevant words as shown in Figure 2.

The same process was followed for neutral reviews also. Tokenization, lemmatization, and topic modeling were done in order to extract the relevant topics and the words related to each topic.

Words from tokenized and lemmatized neutral reviews were “volume”, “bass”, “treble”, “design”, “long”, “use”, “pain”, “ear”, “reliable”, and “buyer.”

In the next step, the reviews wherein the seller name and even the word seller was mentioned were scraped.

The reviews where the word seller was mentioned were considered and the relevant frequent words were found using the same process as above. Then, topic modeling was done in order to find out the topics on which the sellers can concentrate in order to improve their performance.

The frequent words include “delivery,” “brand,” “quality,” “range,” “genuine,” “original,” and “amazing,” as shown in Table 10.

The top 3 relevant topics and corresponding relevant words are shown below in Figure 3.

For the chosen, Cloudbtail was one of the sellers. So, the reviews wherein the name Cloudbtail was mentioned were scraped. Then these reviews were further classified as positive, negative, and neutral. Their corresponding percentages were also found.

Now, these percentages were compared with the Cloudbtail positive, negative percentages, which are calculated in general on the whole for all the deliveries related to various products by the seller. If there is a lot of variation, it is recommended to choose another seller, otherwise it is fine to proceed and order the product from the chosen seller.

There were 110 reviews that mentioned the name of the seller (Cloudbtail). Out of 110, 89 were positive, 16 were negative, and 5 were neutral, as shown in Table 11.

Figure 2.

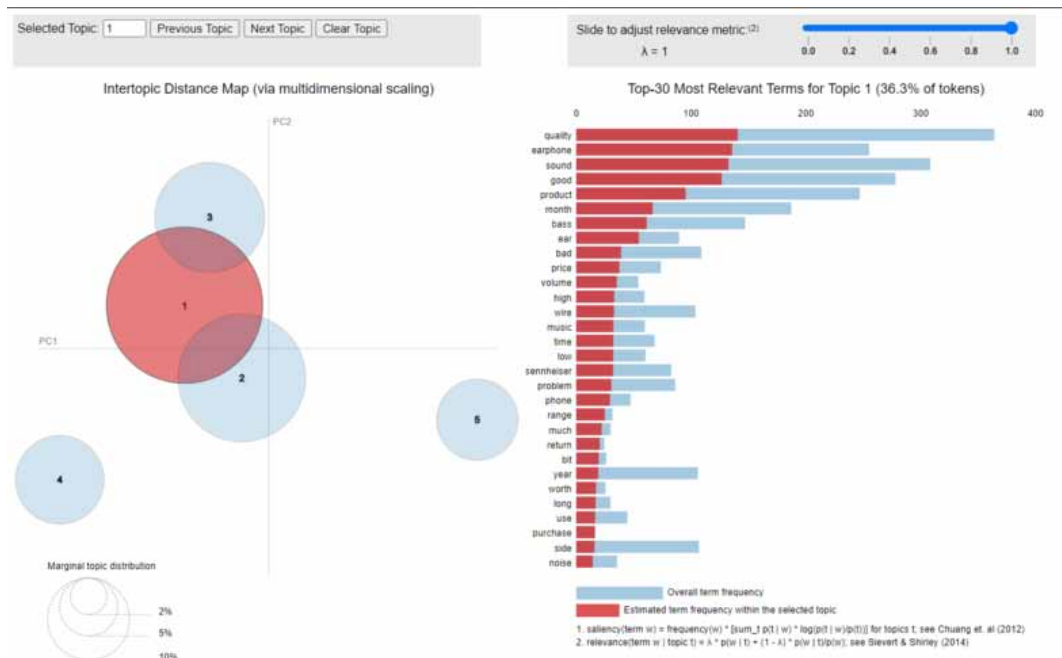


Table 10. Top 30 frequent words in seller reviews after Lemmatization

Word	Count	Word	Count
product	250	amazing	40
seller	200	great	40
good	200	original	35
earphone	170	cloudtail	35
quality	160	delivery	35
sound	140	brand	35
bass	80	awesome	35
sennheiser	70	high	30
price	65	were	30
music	60	month	30
headphone	55	review	30
ear	55	range	30
time	50	genuine	30
year	45	treble	25
warranty	45	day	20

Figure 3.

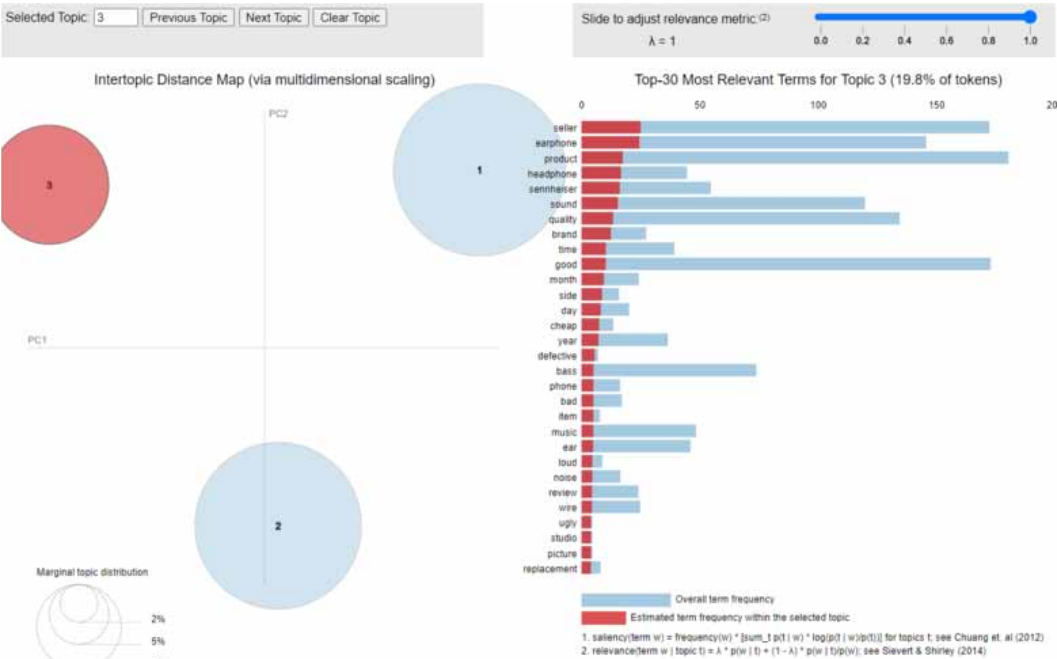


Table 11. Percentage of analyzed seller reviews

Sentiment	No. of Reviews	Percentage
Positive	89	80.91%
Neutral	5	4.55%
Negative	16	14.55%

Comparing the above analysis for a specific product with the Cloudbail India sentiment analysis on all the deliveries as shown in Table 12, it can be observed that our analysis was nearly on the same lines as the general analysis. This indicates that it is fine to choose the seller and the consumer can proceed to purchase the product from that particular seller.

So, scraping the sample product reviews in this manner, categorizing them into positive, negative, neutral, and finding out the relevant frequent words and topics help the visitor to find the pros and cons of the product easily.

Regarding the impact on third-party sellers, in the existing process the visibility of the seller is not on a granular level. The recommendation given by the e-commerce site is based on the overall performance of the seller. So this study helps in providing the results on a granular level and the seller suggestions for the end user are based on the performance of the seller for that specific product.

Scraping the reviews where the word seller is mentioned and finding out the frequent words and topics would help the seller to improve upon certain areas.

Scraping the reviews wherein a particular seller name is mentioned, finding out the positive, negative, and neutral percentages and comparing them with the percentages the seller has, on the whole, would help the consumer to judge the authenticity. This way it would benefit the consumer to choose the seller based on the reviews rather than depending only on the ratings.

The above method, if followed, would help in improving the customer experience with the seller, and help them to take an insightful decision; it is a win-win for both the consumer and seller.

DISCUSSION

The importance of online reviews in consumer decision making has already been well established (Huang and Pape, 2020). Since the last few years, as e-commerce has been on the rise, its value is increasing day by day. Majority of the consumers prefer to read other people's opinion before going forward with the product they have selected. More the product reviews, it is more likely that the consumer will spend more time deciding on that product rather than any of the others, and will likely end up purchasing it (Kunlin et al., 2020).

Many studies have already talked about the impact of online reviews in customer decisions, sentiment analysis on online product reviews in order to find out the opinion of the customer regarding the product, recommendation systems, predicting what the customer would order in the future using machine learning and deep learning algorithms (Dennis et al.,2020). The factor that differentiates

Table 12. Percentage of Seller reviews in General

Sentiment	Lifetime Percentage
Positive	89%
Neutral	4%
Negative	7%

this study from all the above-mentioned ones is that even the seller reviews have been considered for further analysis. The seller reviews usually are present for each seller on a high level, for all the orders that the seller has delivered. This can be used to judge the performance of the seller on a general level. But, when the consumer wants to select a particular seller while ordering a particular product, it is necessary to even present the seller's performance in delivering that particular product. The past delivery experience and the reviews that are given by the consumers for the delivery of that product should also be considered. Though this is on a more granular level, this marks an important aspect for the seller. This aspect of even considering seller reviews, which was not done in previous analyses, helps us to not only get a clear idea about the pros and cons of the seller but also to analyze their performance and decide on the seller. According to Huang and Pape (2020), the literature suggests that two effects of consumer reviews on sales are the awareness effect and the persuasive effect. The authors found the effect of both, but the persuasive effect is slightly larger than the awareness effect. Hence we analyzed the effect of both consumer reviews and seller reviews.

So, briefly, this study mainly reflects the importance of mining online product reviews and also analyzing the impact these reviews create on third-party sellers. This study is beneficial to both the consumer and the seller. Through this study, it is made clear that the seller reviews also carry equal importance as product reviews. The seller reviews do not only mean the reviews on the whole but also the reviews of the seller given by the customers in the product review itself. The seller should also consider them in order to take further decisions on how and in what areas to improve. The techniques used and the method followed can be utilized for various kinds of products too in order to analyze the impact the reviews create on third-party sellers. These results may be true for younger adult consumers but not for older adult consumers (von Helversen et al., 2018). Also, the study done by von Helversen et al. (2018) suggests that the older adult consumers do not consider positive reviews focusing on positive experiences but are easily persuaded by reviews reporting negative experiences.

Though some of the online giant retailers have realized the fact and are making progress toward analyzing these reviews and gaining insights, it is high time that others also realize the same. Apart from the reviews, people also look at the way they are presented. The navigation should be easily understood so that the visitor can find them and should like to go through them. The seller reviews should also be clearly visible to the consumer, in order to avoid those cases where there is an issue with the seller delivery. Though retailers automatically recommend sellers with high ratings and popularity, when the different sellers also catch the attention of the customer, he/she can decide whom to select for the delivery of that particular product.

Managerial Implications

The implications of this study are many. The modeling done in this study would help the consumer to know the positive and negative aspects of a product. The final results would help the consumer to directly identify the pros and cons of not only the product but also the seller. The specific areas where the product excels and those areas where the product lags can also be easily identified using this study. The results would also help the consumer to decide on a particular seller for that product by comparing the performance. In order to improve product sales performance, the seller review analysis would be of great help. Apart from the mails for ratings to the consumer after the product gets delivered, continuous requests should also be sent to the consumer to write their experiences, in form of reviews mentioning the seller name, about the delivery and also about the product after three/four days of usage/delivery. This enables the seller to analyze their past performances and, taking into account those observations, they can improve in future deliveries.

CONCLUSION AND RECOMMENDATIONS

This study mainly reflects the importance of mining online product reviews and also analyzing the impact these reviews create on third-party sellers. This study is beneficial to both the consumer and

the seller. Through this study, it is made clear that the seller reviews also carry equal importance as product reviews. The seller reviews do not only mean the reviews on the whole but also the reviews given by the customers to the seller in the product review itself. The seller should also consider them in order to take further decisions on how and in what areas to improve.

Considering the analysis done in this study, the insights observed/gathered help both the consumer and the seller. The consumer benefits from the fact that instead of going through a lot of reviews in order to know about the pros and cons of the product, this analysis helps him/her to directly view the percentage of positive, negative, and neutral reviews and the relevant frequent words in each category, and also drives home the fact that there are certain topics that the enduser can view directly with the corresponding words in each topic. This would help him to get an idea of the product in less time. The seller comparison on the whole and on a particular product would also help the consumer to decide whether to select that particular seller or choose another.

From the seller's point of view, in order to improve upon product sales performance, the seller review analysis would be of great help.

This study, by considering seller reviews, which was not done in previous analyses, helps us to not only get a clear idea about the pros and cons of the seller but also to analyze the performance and decide on the seller. The techniques used and the method followed can be utilized for various kinds of the product too in order to analyze the impact the reviews create on third-party sellers

Directions for Future Research

Although sentiment analysis was used in this study, it is not limited to it. Many other methods in NLP can be implemented for analyzing and understanding online reviews. Some of them include summarizing reviews into a paragraph or bullet points, identifying which products are popular by extracting entities from reviews, and identifying emerging trends based on the timestamp of the reviews. Speech recognition tools and chatbots can also be created in order to answer the specific questions of customers. This study can also be extended to creating a recommendation model for consumers based on their previous purchases. Finally, though the future looks extremely challenging, there are many advancements in this discipline (NLP and also machine learning in particular), and in the coming years it is very likely that these developments would make complex applications look possible.

Limitations

Though everything seems to be in place and the process seems to be easy, there are certain limitations in this study.

One limitation of this study is transparency. When a consumer purchases a product and writes a review only mentioning the product but not the seller, online retailers like Amazon would have the control to display the seller's name under the title of the review. Though they mark the reviews with a Verified Purchase tag, mentioning the seller name also would benefit the consumer.

Though the recommendation of the seller by Amazon is based on many factors such as ratings, the seller list is not clearly visible and a small link is available to view the other sellers. This causes a huge impact on the other third-party sellers who go unnoticed. Also, many sellers like Cloutail and Appario are owned by Amazon or have a stake in them. These sellers are usually recommended by Amazon for the majority of the products. Though it appears that a giant retailer like Amazon provides a platform wherein there are equal opportunities for third-party sellers, this doesn't seem to be the actual case. Due to all such reasons, people usually do not notice the third-party sellers. Since the purchases are less from these sellers, the ratings and reviews are even lower when compared to the best seller (owned by Amazon).

REFERENCES

- Al-Obeidat, F., Spencer, B., & Al Taii, M. (2017). Identifying major tasks and minor tasks within online reviews. *Future Generation Computer Systems*.
- AL-Sharuee, M. T., Liu, F., & Pratama, M. (2020). Sentiment analysis: Dynamic and temporal clustering of product reviews. *Applied Intelligence*, 1–20.
- Ali, M. M., Doumbouya, M. B., Louge, T., Rai, R., & Karray, M. H. (2020). Ontology-based approach to extract product's design features from online customers' reviews. *Computers in Industry*, 116, 103175. doi:10.1016/j.compind.2019.103175
- Anton, B., & Martin, B. (2020). Using VADER Sentiment and SVM for Predicting Customer Response Sentiment. *Expert Systems with Applications*, 113746.
- Caitlin, D., Theresa, A., Philip, E., & Suzzane, B. (2019). A systematic review of natural language processing and text mining of symptoms from electronic patient-authored text data. *International Journal of Medical Informatics*, 125, 37–46. doi:10.1016/j.ijmedinf.2019.02.008 PMID:30914179
- Camilleri, A. R. (2020). The importance of online reviews depends on when they are presented. *Decision Support Systems*, 133, 113307. doi:10.1016/j.dss.2020.113307
- Chang, Y. C., Ku, C. H., & Chen, C. H. (2019). Social media analytics: Extracting and visualizing Hilton hotel ratings and reviews from TripAdvisor. *International Journal of Information Management*, 48, 263–279. doi:10.1016/j.ijinfomgt.2017.11.001
- Chauhan, U. A., Afzal, M. T., Shahid, A., Abdar, M., Basiri, M. E., & Zhou, X. (2020). A comprehensive analysis of adverb types for mining user sentiments on Amazon product reviews. *World Wide Web (Bussum)*, 1–19. doi:10.1007/s11280-020-00785-z
- Dennis, K., Stefan, L., & Markus, S. (2020). Predicting online shopping behaviour from clickstream data using deep learning. *Expert Systems with Applications*, 113342.
- Doi, N., & Hayakawa, H. (2020). Electronic word-of-mouth: A survey from an economics perspective. *International Journal of the Economics of Business*, 1–18.
- Fan, C., Tianlin, Z., Yuxin, W., & Nan, G. (2016). Research on Classifying Web Information with Latent Semantic Based on Modified LDA Model. *International Conference on Networking and Network Applications*. doi:10.1109/NaNA.2016.25
- Gallagher, C., Furey, E., & Curran, K. (2019). The application of sentiment analysis and text analytics to customer experience reviews to understand what customers are really saying. *International Journal of Data Warehousing and Mining*, 15(4), 21–47. doi:10.4018/IJDWM.2019100102
- Guang, W., Jainfang, J., & Shen, Y. (2017). Quality-related Fault Detection Approaches Based on Data Preprocessing. *IFAC-PapersOnLine*, 50(1), 15740–15747. doi:10.1016/j.ifacol.2017.08.2305
- Hong, L., Hao, T., Wei, X., & Ziyi, G. (2016). Sequential Bag-of-Words model for human action classification. *CAAI Transactions on Intelligence Technology*.
- Huang, M., & Pape, A. D. (2020). The Impact of Online Consumer Reviews on Online Sales: The Case-Based Decision Theory Approach. *Journal of Consumer Policy*, 43(3), 463–490. doi:10.1007/s10603-020-09464-y
- Ifie, K. (2020). Excellent product... but too early to say: Consumer reactions to tentative product reviews. *Journal of Interactive Marketing*, 52, 35–51. doi:10.1016/j.intmar.2020.03.002
- Ike, V., & Sathish, A. (2020). A review of topic modeling methods. *Information Systems*, 101582.
- India Brand Equity Foundation (IBEF). (2020). *Indian Ecommerce Industry Report*. Accessed at <https://www.ibef.org/industry/e-commerce-presentation>
- Jatinderkumar, R., & Rajnish, M. (2016). On continent and script-wise divisions-based statistical measures for stop-words lists of international languages. *Procedia Computer Science*, 89, 313–319. doi:10.1016/j.procs.2016.06.076

- Jiaxiu, H., Xin, W., Mark, B., & Barrie, R. (2020). Revealed preference in online reviews: Purchase verification in the tablet market. *Decision Support Systems*, 113281.
- Khaltar, B. O., & Atsushi, F. (2009). A lemmatization method for Mongolian and its application to indexing for information retrieval. *Information Processing & Management*, 45(4), 438–451. doi:10.1016/j.ipm.2009.01.008
- Kordrostami, E., & Rahmani, V. (2020). Investigating conflicting online review information: Evidence from Amazon. com. *Journal of Retailing and Consumer Services*, 55, 102125. doi:10.1016/j.jretconser.2020.102125
- Kunlin, L., Yuhan, C., & Liyi, Z. (2020). Exploring the influence of online reviews and motivating factors on sales: A meta-analytic study and the moderating role of product category. *Journal of Retailing and Consumer Services*, 55, 102107. doi:10.1016/j.jretconser.2020.102107
- Li, T., Wang, X., & Wu, Y. (2020). Pricing strategies in presence of online consumer ratings-from the product customization perspective. *Journal of Theoretical and Applied Electronic Commerce Research*, 15(3), 84–100. doi:10.4067/S0718-18762020000300107
- Lusiana, C., & Meilianaand Alvin, C. (2019). Social Media Web Scraping using Social Media Developers API and Regex. *Procedia Computer Science*, 157, 444–449. doi:10.1016/j.procs.2019.08.237
- Mahmoud, E., Paulo, A., Paul, R., Martin, W., & Steven, Y. (2019). Retrieving, classifying and analysing narrative commentary in unstructured (glossy) annual reports published as PDF files. *Accounting and Business Research*, 50(1), 6–34.
- Obeidat, Z. M. I., Xiao, S. H., Iyer, G. R., & Nicholson, M. (2017). Consumer revenge using the internet and social media: An examination of the role of service failure types and cognitive appraisal processes. *Psychology and Marketing*, 34(4), 496–515. doi:10.1002/mar.21002
- Qing, X., Xinyuan, Z., Ying, D., & Min, S. (2020). Monolingual and multilingual topic analysis using LDA and BERT embeddings. *Journal of Informetrics*, 14(3), 101055. doi:10.1016/j.joi.2020.101055
- Ramaswamy, S., & DeClerck, N. (2018). Customer perception analysis using deep learning and NLP. *Procedia Computer Science*, 140, 170–178. doi:10.1016/j.procs.2018.10.326
- Sameh, A., & Ozgur, T. (2020). A comparative assessment of sentiment analysis and star ratings for consumer reviews. *International Journal of Information Management*, 54, 102132. doi:10.1016/j.ijinfomgt.2020.102132
- Sann, R., & Lai, P. C. (2020). Understanding homophily of service failure within the hotel guest cycle: Applying NLP-aspect-based sentiment analysis to the hospitality industry. *International Journal of Hospitality Management*, 91, 102678. doi:10.1016/j.ijhm.2020.102678
- Sasikala, P., & Sheela, L. M. I. (2020). Sentiment analysis of online product reviews using DLMNN and future prediction of online product using IANFIS. *Journal of Big Data*, 7(1), 1–20. doi:10.1186/s40537-020-00308-7
- Shobana, J., & Murali, M. (2020). A survey on sentiment analysis algorithms and techniques. *Journal of Advanced Research in Dynamic and Control Systems*, 1893–1906.
- Vera, S., Yiftach, B., Eli, K., & Eyal, K. (2020). Deep Learning for Natural Language Processing in Radiology—Fundamentals and a Systematic Review. *Journal of the American College of Radiology*. PMID:32004480
- Viji, D., Asawa, N., & Burreja, T. (2020). Fake reviews of customer detection using machine learning models. *International Journal of Advanced Science and Technology*, 86–94.
- Vineet, K., Arpita, K., Rosy, B., & Marta, B. C. (2020). Why do online retailers succeed? The identification and prioritization of success factors for Indian fashion retailers. *Electronic Commerce Research and Applications*, 39, 100906. doi:10.1016/j.elerap.2019.100906
- von Helversen, B., Abramczuk, K., Kopeć, W., & Nielek, R. (2018). Influence of consumer reviews on online purchasing decisions in older and younger adults. *Decision Support Systems*, 113, 1–10. doi:10.1016/j.dss.2018.05.006
- Wang, Z., & Chen, Q. (2020). Monitoring online reviews for reputation fraud campaigns. *Knowledge-Based Systems*, 195, 105685. doi:10.1016/j.knosys.2020.105685

Wen, S., Wenli, L., & Shidao, G. (2020). Effect of online product reviews on third parties' selling on retail platforms. *Electronic Commerce Research and Applications*, 39, 100900. doi:10.1016/j.elerap.2019.100900

Xianhao, X., Yaohan, S., Wanying, C., Yeming, G., & Hongwei, W. (2020). Data-driven decision and analytics of collection and delivery point location problems for online retailers. *Omega*, 102280.

Xiaolin, L., Chaojiang, W., & Feng, M. (2019). The effect of online reviews on product sales: A joint sentiment-topic analysis. *Information & Management*, 56(2), 172–184. doi:10.1016/j.im.2018.04.007

Xiong, L., Xiaodong, Z., Wangtu, X., & Wei, P. (2020). Measuring ease of use of mobile applications in e-commerce retailing from the perspective of consumer online shopping behaviour patterns. *Journal of Retailing and Consumer Services*, 55, 102093. doi:10.1016/j.jretconser.2020.102093

APPENDIX

Python Code

```
import pandas as pd
import nltk
import matplotlib as plt
import seaborn as sns
reviews = pd.read_csv("Seller Reviews.csv")
reviews
reviews.describe()
reviews.info()
content = reviews['content']
content
l = len(content)
l
df = []
from nltk.sentiment.vader import SentimentIntensityAnalyzer
sid = SentimentIntensityAnalyzer()
for i in range(l):
    scores = sid.polarity_scores(content[i])
df.append([scores])
df2 = pd.DataFrame(df, columns=['Score'])
df2.to_csv('Cloudtail sentimentscore.csv')
word2count = {}
for data in range(l):
    words = nltk.word_tokenize(content[data])
    for word in words:
        if word not in word2count.keys():
            word2count[word] = 1
        else:
            word2count[word] += 1
print(word2count)
import heapq
freq_words = heapq.nlargest(200, word2count, key=word2count.get)
print(freq_words)
ar = pd.read_csv("AmazonReviews.csv")
arp = ar['Compound Sentiment']
#Topic Modeling
from nltk import FreqDist
pd.set_option("display.max_colwidth", 200)
import numpy as np
import re
import spacy
pip install gensim
import gensim
from gensim import corpora
pip install pyLDAvis
import pyLDAvis
import pyLDAvis.gensim
import matplotlib.pyplot as plt
```

```
pos = pd.read_csv("Positive Reviews.csv")
pos.head()
neg = pd.read_csv("Negative Reviews.csv")
sel = pd.read_csv("Seller Reviews.csv")
cld = pd.read_csv("Cloudtail Reviews.csv")
#Frequent Words
deffreq_words(x, terms = 30):
all_words = ' '.join([text for text in x])
all_words = all_words.split()
fdist = FreqDist(all_words)
words_df = pd.DataFrame({'word':list(fdist.keys()),
'count':list(fdist.values())})
    # Selecting top 20 most frequent words
    d = words_df.nlargest(columns="count", n = terms)
plt.figure(figsize=(20,5))
    ax = sns.barplot(data=d, x= "word", y = "count")
ax.set(ylabel = 'Count')
plt.show()
#Positive Topic Modeling
freq_words(pos['content'])
freq_words(neg['content'])
freq_words(sel['content'])
sel['content'] = sel['content'].str.replace("[^a-zA-Z#]", " ")
from nltk.corpus import stopwords
stop_words = stopwords.words('english')
# Function to remove stopwords
defremove_stopwords(rev):
rev_new = " ".join([i for i in rev if i not in stop_words])
    return rev_new
# Remove short words (length < 3)
sel['content'] = sel['content'].apply(lambda x: ' '.join([w for w
in x.split() if len(w)>2]))
# Remove stopwords from the text
reviews = [remove_stopwords(r.split()) for r in sel['content']]
# Make entire text lowercase
reviews = [r.lower() for r in reviews]
freq_words(reviews, 30)
import spacy
import en_core_web_sm
nlp = spacy.load('en_core_web_sm', disable=['parser', 'ner'])
def lemmatization(texts, tags=['NOUN', 'ADJ']): # filter noun and
adjective
    output = []
    for sent in texts:
        doc = nlp(" ".join(sent))
output.append([token.lemma_ for token in doc if token.pos_ in
tags])
    return output
tokenized_reviews = pd.Series(reviews).apply(lambda x: x.split())
print(tokenized_reviews[1])
```

```
reviews_2 = lemmatization(tokenized_reviews)
print(reviews_2[1])
reviews_3 = []
for i in range(len(reviews_2)):
    reviews_3.append(' '.join(reviews_2[i]))
sel['content'] = reviews_3
freq_words(sel['content'], 30)
dictionary = corpora.Dictionary(reviews_2)
doc_term_matrix = [dictionary.doc2bow(rev) for rev in reviews_2]
LDA = gensim.models.ldamodel.LdaModel
# Build LDA model
lda_model = LDA(corpus=doc_term_matrix, id2word=dictionary, num_
topics=3, random_state=100, chunksize=1000, passes=50)
lda_model.print_topics()
pyLDAvis.enable_notebook()
vis = pyLDAvis.gensim.prepare(lda_model, doc_term_matrix,
dictionary)
pyLDAvis.show(vis)
```

Akash Phaniteja is an MBA student specializing in Business Analytics from SCMHRD, Pune. He comes with an experience of nearly three years having worked as a Pre-Sales Executive. His interest lies in Machine Learning, Deep Learning and Natural Language Processing.