

Decoding user interaction dynamics on Facebook fan page: A sentiment mining approach

Shabana Chandrasekaran^{a,*}

Supriya Kumar De^a

^aXavier School of Management (XLRI), Circuit House Area, Jamshedpur, India 831001

* Corresponding Author

Abstract

Increased presence of firms on social media like Facebook, has enabled users to interact directly with the firms. As this platform is used for brand promotions, it is vital for firms to understand the dynamics of the user interaction that happens on the brand's fan page. This paper investigates the influence of marketer generated contents and the recursive effect of sentiments on user interaction with the help of sentiment mining approach and Negative Binomial Regression. The analysis is done over 2121 marketer contents and 22 million user contents across 5 Facebook brand pages. The results show that increase in positive comment sentiment triggers increased liking and sharing behaviour by the users on the marketers' content. And users are more inclined to like and share on marketers' contents that are informative posts. Posts containing images garner higher likes and comments but videos have higher shares. The results are useful for the brands to understand user behaviour on a brand's fan page.

Keywords:

User interaction, Sentiment mining, Social media, Facebook fan page

1. Introduction

Web 2.0 has enabled users to create and share high volume content about products and services which firms now prefer to listen to and mold the business activities for projecting the brand positively. The proliferation of applications, in terms of blogs, social media etc., on Web 2.0 have driven content sharing to a new dimension where consumers look forward to other consumers to know about a product or a service than looking forward to the marketers of the product. This has turned the classical one-way marketing communication of brands to a communication that entrusts customers to choose their own spokesperson for a brand (Del Giudice, Della Peruta & Carayannis, 2013). Realizing the need for more innovative and authentic channels of marketing communication, marketers look forward to digital media. Among digital media forums, social media has generated deeper impact as the shared content generates empathy among the users, thereby acting as a platform for 'emotional contagion' (Hatfield, Cacioppo, & Rapson, 1993). The emotions influence the level of enthusiasm and directs the way in which interaction happens on the platform (Bickart & Schindler, 2001). We can generally observe that when a user posts positively about a brand, others follow the discussion and the discussion is more likely to be observed by others. This active participation of consumers on social media has shifted the 'power-balance' to the customer side (Van Dijk, 2006). Consumers' purchase decisions are now powered by product reviews generated by like-minded consumers than through advertisements and other form of personal communication (Dellarocas, 2006). As social media has become a critical channel for brand-related information, it is of utmost importance for firms to understand the various aspects of the communication that would drive the user interaction.

Researches have focused on understanding the consumer interaction in social media platforms. Using netnographic analysis the behavior of online communities were studied to identify the consumer decision making process (De Valck et al., 2005). The importance of consumer to consumer interactions on social media as a critical element of marketing promotional mix was explored by Mangold & Faulds (2009). Conceptual framework has been developed based on consumer participation and motivation derived out of the consumer activities in social media (Heinonen, 2011). Research has also been done to capture the dynamic and fluctuating nature of user interactions on social media (Brodie et al., 2013). However, these researches have isolated consumer interactions that are eventually the reactions to a marketer communication. As social media is a platform where consumers interact with both the firm and other consumers, and user reactions are responses to marketers content on social media, it is vital for marketers to understand the interaction with the consumers. It is therefore of significance to study the impact of user sentiments on user interaction and the influence of marketers' communication on user interaction, as this will aid marketer to understand the dynamics of user interaction in social media.

A few studies have explored the marketer-user interaction on social media, in various dimensions especially on Facebook. User interactions were studied in terms of brand posts' popularity of Facebook fan page (De Vries et al., 2012; Khobzi & Teimourpour, 2014 and Sabate et al., 2014), online user engagement (Luarn, Lin & Chiu, 2015), persuasiveness in repeated purchases (Goh, Heng & Lin, 2013) and as a metric towards brand building (Moro, Rita, & Vala, 2016). However, only a few studies have attempted to implement a model that would effectively capture the impact of user content on the user interaction. User interaction in a brand's fan page is an 'iterative engagement process' (Hollebeek, 2011) and is governed by the emotions embedded in

the shared post. The emotions are the varying sentiments of the users which influences information processing and interpretation by the users (Salehan and Kim, 2016) thereby impacting consequent user interaction. Hence the model should also capture the influence of existing sentiments on subsequent interaction. This will enable marketers to secure increased control on the information dissemination process. Text mining provides an interesting approach for elucidating knowledge from raw data. Its application in Facebook has been studied as a part of social media monitoring (Cvijikj & Michahelles, 2011 September), to understand the impact of sentiments on brand awareness (Zhang et al., 2014, September) and target selection for online advertising (Zhang et al., 2016). Akin to these studies, we apply sentiment mining to understand user interaction that enables marketers to design an appropriate marketing communication.

While looking at the influence of contents shared by the marketer on users in Facebook from a marketer's perspective, it is crucial to understand the reactions and responses of user participation and contribution in online brand communities. Consumer to consumer communication, mediated by social media have loosened the control that firms held on information dissemination (Mangold, & Faulds, 2009). Therefore to regain control, firms are keen to understand not only the user dynamics but also the responses to the marketing communication in the brand communities. Some studies have investigated the impact of marketers' content on user reactions (Cvijikj, Spiegler & Michahelles, October, 2011; Shen & Bissell, 2013; Sabate et al., 2014 and Tafesse, 2015). Since we observe that the findings from the studies were varying in nature about the influential effects of marketer communication, we revisit the existing claims to understand what type of Facebook fan page posts shared by the marketers attracts more user interaction.

For our study, we collect data from the top 5 Facebook brand pages¹ from India that has the highest fan counts for the year 2016. While capturing the user sentiments of Facebook data from India, we observe predominant usage of Hindi texts in Roman scripts. The existing sentiment mining techniques is not customized to effectively handle such texts. Therefore, our study addresses these drawbacks by applying a sentiment mining approach that overcomes the challenge of Hindi texts in Roman scripts. Social media data rarely displays an equal variance as the interaction level need not represent homogeneity across samples. So we require an appropriate model to capture user interaction that would also accommodate the non-homogeneity. Unlike previous studies on social media data that have used Ordinary Least Square models (Sabate et al., 2014 and Khobzi & Teimourpour, 2014) and Poisson model (De Vries et al., 2012) to model user interaction, our study has utilized Negative Binomial Regression model in explaining the same. Negative Binomial Regression is appropriate for modelling count data and also relaxes the assumptions of variance that has to be strictly adhered for other models (Hilbe, 2011) and it has been used earlier by Salehan & Kim (2016) to model readership measured using upvotes.

Though the prime objective is to study the impact of user sentiments, we begin by examining the influence of marketers' content as it initiates user interaction which subsequently gets driven by the shared user sentiments. We first look into the characteristics of the post shared by the marketer in a brands fan page. We look at two characteristics of posts namely content type and media type. As different characteristics attract different levels of user interaction, we study the kind of post content and media content that triggers higher user interaction. Next we review the

¹ <https://www.socialbakers.com/statistics/facebook/pages/total/india/brands/>

influence of existing user sentiment on subsequent interaction level, which is addressed by extracting the user sentiments using sentiment mining technique and modelling user interaction.

The rest of the paper is organized as follows: Section 2 presents the conceptual framework that describes the user interaction process. In section 3, we describe data collection and data analysis pertaining to 2121 posts. Results of the analysis are discussed in section 4. This is followed by implications of the study in section 5 and conclusion in section 6.

2. Conceptual framework and hypotheses

The rise of social media has shifted the focus from the contents that are shared by the marketer alone to the contents shared by both the marketer and the user. The users are not only potential consumers but also co-creators, partnering with the firm to mold the product offering as per needs. Therefore it is of significant importance for the marketers to engage and maintain a healthy user interaction amongst the users as it provides valuable insights. Here, the marketer contents plays a pivotal role in directing the user engagement on the social media platform and marketers use various techniques to improve user interaction. In Facebook, marketers share various contents or posts about the product or service on the brand fan page. The posts shared are designed in a way that they are attractive enough to trigger user interaction on the respective fan page. These posts constitute the Marketer Generated Contents (MGC). The crucial features that define the fan page post is its content type and the media type (Cvijikj, Spiegler & Michahelles, October, 2011). The fans of a brand, essentially the users who have liked the brand's fan page, react to the posts by liking, commenting and sharing. These metrics are considered as the User Generated Contents (UGC) on the Facebook platform. While liking and sharing a post are objective in indicating the preference towards a fan page post, commenting on a post is quite subjective as the user may comment positively or negatively. Hence it is crucial to classify the

comments according to the sentiments- positive, negative or neutral because positive posts may increase the liking and sharing behavior. With the available data from the MGC and UGC, we study the effect of various features of fan page post on the user interaction that results through likes, shares and comments. Further, when a user comments positively for brand's fan page post, the user may also show affinity towards it by liking and sharing the fan page post. On the other hand a user posting a negative comment on a brand's fan page post is less likely to like or share the post. Hence, we also study the influence of user sentiments on the user preference to like and share an MGC.

In this study, we examine the significance of contents shared by the brands on its Facebook fan page and its impact on user interaction, measured through the number of Likes, Shares and Comments. We further explore the effect of user sentiments on the metrics of user interaction across brands through sentiment mining approach. Fig. 1 shows the conceptual framework.

2.1. Studies on MGC

2.1.1. Content type

The earliest study to identify information and entertainment as the content categories of post in spite of being insignificant in explaining the user interaction was done by De Vries, Gensler, & Leeflang (2012). However, these content categories formed the reference to classify fan page post and was later extended by other studies to accommodate more categories. Shen & Bissell (2013) classifies post contents as events, product-related, promotion-related, entertainment and others and finds that entertainment attracted more number of comments. Another interesting classification of post contents was with respect to the product, price, place and promotional information (Thongmak, 2015, June). The study found that promotional posts significantly

attracted more comments than likes and shares. A wide number of other studies categorized post content as informational, entertainment-oriented, remunerative and social (Goh, Heng & Lin, 2013; Sabate et al, 2014 and Luarn, Lin & Chiu, 2015). Based on this, we classify the posts as Informative- posts that communicate information about the brand or product (Muntinga et al., 2011), Social- posts that prompt the users to respond (eg: share your views) (Cvijikj & Michahelles 2013), Entertainment- posts that broadcasts non-brand or non-product related information that enables hedonic pleasure among users (eg: comedies, movies and sports related information) , and Remuneration- posts that communicates special offers, discounts and contests (Wood et al., 2013). Based on these classification, we observe that social and remunerative posts intend to increase user interaction, by inducing user reactions to the post, over other types of post. To validate our assumption, we propose:

a.) User interaction on a Facebook post varies with content type across brands of a product category

H1: The number of likes differs across the content types among brands

H2: The number of comments differs across the content types among brands

H3: The number of shares differs across the content types among brands

2.1.2. Media type

Expressions take various forms of cues like text, image or video. In an online medium, contents are appealing when users perceive a post to be more expressive (Steuer, 1992). Online contents are expressed through various cues like links that diverts attention to brand's website, status, photos and video that gives information about the brand or the product. The extent to which a user finds these cues stimulating is defined as vividness (De Vries et al., 2012). A video has the

highest level of vividness followed by links (as they direct users to websites that has more vivid media types), images and status (Luarn, Lin & Chiu, 2015). The significant effect of brand's post vividness on user interaction has been observed in many studies and have reported varying results. While one study shows that links and videos garnered high user interaction (De Vries et al., 2012) other has observed that photos attracted higher user interaction (Sabate et al., 2014). Studies have also shown that status post, that has least vividness has significantly high user interaction (Cvijikj, Spiegler & Michahelles, October, 2011 and Moro, Rita, & Vala, 2016) and also has no significance of the effect of media type on user interaction (Tafesse, 2015). Hence we would like to study the performance of various media types. Deriving from the above studies we propose

b.) User interaction on a Facebook post increases with increasing vividness of media type across brands

H4: The number of likes increases with increasing vividness of media type across brands

H5: The number of comments increases with increasing vividness of media type across brands

H6: The number of shares increases with increasing vividness of media type across brands

2.2. Studies on user sentiment

The research on the customers' attitude towards a page's post using the sentiment of the comments considers the popularity of the post through number of likes and comments. The sentiment strength of the comments that ranges from negative to positive is also found to influence subsequent user interaction. The earliest study to utilize user sentiments in understanding user interaction on Facebook was by Cvijikj & Michahelles, (2011, September). The study identifies that positive sentiments promoted more post shares. This was subsequently

tested by De Vries et al., (2012) and results suggested that positive user sentiments were significant for increased number of likes and comments. Khobzi & Teimourpour (2014) also modelled the impact of sentiments on the number of shares and likes and found that user sentiment was significant in increasing the number of likes and number of comments. As users talk positively about a post, the post is bound to be liked and shared more. Hence we posit:

H7: Sentiment of a post positively influences the Like count across brands

H8: Sentiment of a post positively influences the Share count across brands

2.3. Studies on sentiment analysis

Analyzing sentiments is a complex process, as the meaning of words tend to change according to the context and in many situations the accuracy of an analysis tool is influenced by the context in which the words are used (Turney, 2002 and Kang, Yoo & Han, 2012).

As sentiment mining is a vast area with custom developed techniques suitable for individual application, we go through some of the sentiment mining techniques that have been developed for Facebook platform. Researches have used supervised techniques, unsupervised techniques or a combination of both to classify sentiments. Supervised techniques use machine learning algorithms which rely on labelled training data sets for improved classification accuracy, while unsupervised techniques use lexicons or dictionaries of opinion words with pre-determined sentiment value for classification. Many studies that use supervised approach have focussed on the application of different algorithms to capture the sentiment. Sentiment classification of Tunisian Facebook statuses were approached using a combination of support vector machine and Naïve Bayes (Hamouda & Akaichi, 2013). Another similar study on Czech Facebook statuses was performed using maximum entropy classifier (Habernal, Ptáček & Steinberger, 2014).

Meire, Ballings & Van den Poel (2016) predicted Facebook user sentiment with additional information like user profile features, user's previous status mood and other interaction behaviours using Random Forest algorithm. Another study approaches sentiment classification using Naïve Bayes to classify the Facebook status updates uses 2000 training set to achieve better prediction (Troussas et al., 2013). Research has also analysed Facebook sentiments over Hadoop with MapReduce platform using sentiment analysis package of R (Dasgupta et al., 2015). On the hybrid approaches, Ortigosa, Martín & Carro (2014) developed a hybrid of Lexical approach and machine-learning techniques to classify Facebook statuses in an e-learning system to suggest learning activities based on the emotions of the students. Comparative performances of two different fan pages used a combination of supervised and unsupervised clustering to classify the polarity of the comments (Neri et al., 2012). In unsupervised approach, Ngoc & Yoo (2014) custom developed a sentiment lexicon to rank Facebook fan pages using the sentiments of the page users. Product recommendations based on sentiment of the user have been developed using a lexicon of opinion words that defined the sentiment category (Pham et al., 2016). Domain specific sentiment lexicons have been developed by extending existing lexicons to accommodate context relevant words (Weichselbraun, Gindl & Scharl, 2011).

Although we find all the three techniques to be used extensively, lexical approaches have gained preference over other techniques due to its adaptability to the usage context (Jurek, Mulvenna & Bi, 2015) and the advantage of not requiring extensive training data for classification accuracy. Also, as Facebook data deals with very short informal text, lexicon based sentiment mining is recommended over supervised approaches (Tang, Fang & Wang, 2014). So, to accommodate regional languages and lingo and to forego the need for labelled data, we use Lexicon based sentiment mining technique for our study.

3. Method

3.1. Data collection

The brands selected is a combination of mobile service providers, an operating system, an FMCG brand, and an electronics brand. We extract the data from the fan page using Netvizz application (Rieder, 2013). The application is free of legal uncertainties as it extracts information that are available with users' permission for public access. The extracted data contains both MGC and UGC. The MGC is characterised by the posts that marketers share on the fan page. The posts were categorized as Information, Entertainment, Remuneration and Social posts based on their content type using key-word match method. For example, posts that contained keywords like offers, discounts etc., were categorised as informative posts while keywords like wishing, happy were marked as social. Similar procedure was followed for the other two type of categories. This classification though not mutually exclusive, tries best to classify a post in one of the categories. The media type of the post is predefined. The UGC is characterized by the reactions of users to the marketer's post in terms of Likes, Shares and comments. The details of comments and other information shared by the users for a fan page post is extracted from the UGC. Table 1 gives the descriptive statistics about the total number of posts shared by the brand and the number of likes, comments and shares that has been generated for the year 2016. A total of 2121 MGC in the form of Facebook posts that have garnered nearly 21 million likes, 0.18 million shares and 0.38 million comments were retrieved. Out of this, 0.14 million comments were retained for further analysis after data cleaning.

Insert Table 1

3.2. Analysis

The analysis is a two-stage process. The first stage consists of evaluating the sentiment of the comments that are generated by the users (Fig. 2.), which is then followed by hypotheses testing and model building.

The sentiments of the comments are analysed using SentiStrength, a Lexical-based approach for sentiment classification (Thelwall et al., 2010). The wordlist of SentiStrength lexicon has a list of possible affect words, emoticons and even abbreviated word list with their meaning along with their sentiment strength. The sentiment strength of each word defined in the wordlist is assigned a score on a scale of -5 to +5, where -5 indicates extreme negative sentiment and +5 indicates extreme positive sentiment of each word. Select part of speech like nouns, verbs etc., and words that are not in the SentiStrength wordlist are assigned a value of 0. The major advantage of SentiStrength is that, it accepts low quality data and allows the addition of a new wordlist (Thelwall, 2013). However, we faced two limitations while analysing the sentiment of the data collected. The first limitation is that the wordlist when used as such in its default setting, did not reflect the sentiments of the text convincingly. Because the actual SentiStrength scores for majority of the words in the wordlist were between -3 to +3, the sentiment scores were skewed more towards neutral sentiment than the extremities. The second limitation we observed is that, as we are dealing with emerging economy, India, the Facebook comments had prominent Hindi texts in Roman script. Both these limitations were addressed by modifying the actual wordlist and sentiment scores of SentiStrength to reflect the sentiments accurately. Potential words that can reflect the sentiments contextually are chosen from a list that has words ranked according to their frequency of occurrence. The list is generated once the actual comments are cleaned of the stop-words and stemmed. Words that occur 50 times or more in the collection of comments were selected. If a selected word already exists in the wordlist, its sentiment score is inspected and the

sentiment score of the word is scaled by a value of 2. If the word does not exist in the wordlist, then it is added to the existing wordlist with a score value 5 or 4 based on its presence in the first half or second half of the ranking. This method helped to alleviate the issue of presence of Hindi texts as the most prominent Hindi words were also included in the wordlist.

The data was analysed using the modified wordlist to calculate the strength of sentiment of each comment. The comment sentiments were then aggregated for each post and the overall sentiment score of the post was calculated using (Yang, Yang & Wu, 2016).

$$(\sum \text{Positive sentiments} + \sum \text{Negative sentiments}) / (\sum \text{Positive sentiments} - \sum \text{Negative sentiments})$$

The second stage of analysis focusses on testing of hypotheses with number of likes, comments and shares along with content of the post (Entertainment, Information, Social, and Remuneration) and type of post (link, status, photo, video, and event) and model development to measure the user interaction with the calculated user sentiment.

The number of likes and shares are count data and showed that the data was heavily skewed in spite of removing the outliers. So the data was transformed (Moore et al., 1989) to its inverse natural log form and tested for the proposed hypotheses. Further, as we could not observe equal variances across the groups, we are limited with conducting one-way ANOVA tests using Games- Howell post-hoc calculations for the post content type and post media type across the UGC- number of likes, comments and shares, and manually interpret them with the help of chi-square analysis for the interaction effect of post content type and post media type across brands.

4. Results & Discussion

On testing the hypotheses for user interaction across various post contents, we find H1, H2, and H3 are significant at $p < .10$ and there exists a significant difference between informative posts and other type of post contents in terms of likes and shares count. Further, analysis of user interaction across content types shows that posts that focusses on informative contents has garnered the maximum number of likes and shares, while content type 'Social' has garnered the least number of likes, comments and shares (Fig 3.a., 3.b., & 3.c.). Interestingly, remunerative posts are found to perform better next only to informative posts in attracting more number of shares and comments. This implies that informative contents trigger distinct responses as compared to other types of post contents. However the marketers trigger to make the audience act through the socializing content has not gained as many number of likes, comments and shares as other post contents. Thus showing that the user engagement varies for different post content.

Insert Figure 3.a., 3.b. & 3.c.

Analysing the types of post contents shared by the 5 brands, we can infer that each brand exhibits unique digital marketing strategy by sharing different proportions of post focussing on information, entertainment, remuneration and social contents (Table 2). For instance, majority of the brands primarily share informative posts as informative posts share details about the latest offerings of a brand. However, remunerative posts were still preferred as it held a major portion in one of the brands (57% by Brand 4) while such posts were least preferred by other brands. We can also observe that brands preferred sharing informative contents, which is followed by social contents that any other content types.

Insert Table 2

The one-way ANOVA that tests the difference in means of likes, shares and comments (H4, H5 & H6) across media types of the post was found to be significant. However, contrary to expectation that high vividness will trigger higher level of user interaction, we observed that posts sharing a photo that corresponds to lower levels of vividness has the highest level of engagement in terms of likes and comments (Fig 4.a. & 4.b.). While videos categorized as high vividness has the highest share count, followed by photos as expected (Fig 4.c.).

Insert Figure 4.a., 4.b., 4.c.

From a brand-wise analysis across media types we observe again that each brand exhibits different proportions of post' shared as link, status, photo and video (Table 3). We observe that photo is the preferred form of media type among majority of the brands, except for Brand 4 that shares 51% of posts in video format, while status update is least preferred media type. Posts containing links or videos are interchangeably preferred form of media content after photo.

Insert Table 3

Now looking at the brand-wise performance, we observe that Brand 5 has gained the maximum number of likes per post as compared to Brand 1, Brand 2 and Brand 3 (Table 4). Posts of Brand 4 has been shared maximum times in comparison to Brand 1, Brand 3 and Brand 5. Similarly Brand 5 has maximum number of comments in comparison to other brands while Brand 3 being the least. From Table 4, it can be observed that Brand 5 has higher levels of user interaction by sharing 53% informative posts and around 29% remunerative posts while. Especially, more than 90% of Informative and Remunerative posts were shared as photos. Brand 3 that focuses more on informative contents (76%) without any remunerative posts has acquired the least engagement.

Insert Table 4

Brand 4 which shares the highest number of videos has the highest share count. Thus it is observed that the brand that has the highest like and comment count posts informative and remunerative contents as photos. While the brand that has the highest share count posts remunerative contents in video format.

The influence of sentiments on user interaction is modelled using negative binomial regression. We tested the following models for the impact of sentiment on number of likes (equation 1) and number of comments (equation 2).

$$\text{Log (like counts)} = \beta_0 + \beta_1 (\text{sentiment score}) + \beta_2 (\text{Brand 1}) + \beta_3 (\text{Brand2}) + \beta_4 (\text{Brand 3}) + \beta_5 (\text{Brand 4}) + \beta_6 (\text{Brand 5}) \quad (1)$$

$$\text{Log (share counts)} = \alpha_0 + \alpha_1 (\text{sentiment score}) + \alpha_2 (\text{Brand 1}) + \alpha_3 (\text{Brand2}) + \alpha_4 (\text{Brand 3}) + \alpha_5 (\text{Brand 4}) + \alpha_6 (\text{Brand 5}) \quad (2)$$

The proposed model showed a significant influence of sentiments on the number of likes and shares across brands as the goodness of fit of the model over the given explanatory variable sentiment score is significant at $p < .05$ (Table 5) across brands. From the parameter estimates table (Table 6) we can observe that for every unit increase in sentiment score the log count of number of likes changes by 0.83 and the log count of number of shares changes by 0.79. Thus showing significant influence of comment sentiment on user engagement.

Insert Table 5

Insert Table 6

Further comparing the incident rates (Table 6) at which the like counts occur, we can observe that Brand 5 has the highest expected number of like counts followed by Brand 2, Brand 4, Brand 3, and Brand 1. Similarly, the expected number of share counts is highest for Brand 4 followed by Brand 2, Brand 5, Brand 1, and Brand 3. The expected percent change in the

incident rate of the number of likes and number of shares is around 2.3% increase for every unit increase in the sentiment score. Since the unit of a sentiment can be any decimal value between -1 and +1, we hold the sentiment value at an average value of 0.17 to observe its effect across the number of likes and shares across brands (Table 7). We further extrapolate the sentiment values to its extremes (-1 and +1) to calculate the expected number of likes and shares and the values are as given in the table (Table 8) which supports the hypothesis that user interaction increases with increasing user sentiments

Insert Table 7

Insert Table 8

The brand wise measures at this sentiment score shows that Brand 5 gets the maximum like count of 18063 while Brand 1 gets the least count of 914. It is also observed that the average like count increases from 1473 to 7749 for the shift in sentiment score from -1 to +1. Similarly for share count, at a sentiment score of 0.17, Brand 4 is shared 131 times as against Brand 3 with only 10 shares on an average. Thus exhibiting a wide difference in the level of interaction among brands. Further the share count increases from 20 to 97 for the shift in sentiment score from -1 to +1. Thus the positive influence of sentiment on engagement measures is observed in our data set.

5. Implications

The study throws light on dynamics of user behaviour by exploring the underlying mechanism in which user interaction happen in Facebook. We show that the positive emotions expressed by users enter a virtuous cycle that will trigger more positive user reactions. This provides a valuable insight for marketers to manage a healthy, positive user interaction that will further reinforce itself, essentially showing that emotions magnify user interaction. Our findings act as a guide for the marketers regarding the type of content and media type that a brand may share on

its fan page. For example, if the objective is to garner higher number of likes, the marketer can share brand posts with low vividness. The findings also aid marketers to determine the content type that will garner higher user interaction and at the same time prevent marketers to avoid mistakes. For example, informational posts do exhibit increased interaction but in the absence of remuneration posts, there will not be an increased interaction. Also, contrary to our assumptions that social and remunerative posts should garner more user interaction as the primary objective of such posts is to induce user interaction, we observe that users' prefer informative contents than social and remunerative contents. Thus our findings act as a guideline for marketers to promote increased interaction in Facebook fan page.

6. Conclusions

Realizing social media sites as a potential influencer of consumer decision making, firms have begun utilizing social media pages as one of the prime channels for interacting with consumers. So it is essential for firms to understand the user interaction that happens on its brand fan page. Consumers interact in Facebook through Like, Comment and Share. On the practical front, these interactions are a source of valuable insights for the firm about the brand. Our study results support firms to understand how users react to different marketer content. To extract sentiment from the user comment, sentiment mining technique has been used. Social media data is not always homogenous in nature. This paper uses Negative binomial regression to deal with this kind of non-homogeneous data.

References:

1. Bickart, B., & Schindler, R. M. (2001). Internet forums as influential sources of consumer information. *Journal of interactive marketing*, 15(3), 31-40.
2. Brodie, R. J., Ilic, A., Juric, B., & Hollebeek, L. (2013). Consumer engagement in a virtual brand community: An exploratory analysis. *Journal of Business Research*, 66(1), 105-114.
3. Cvijikj, I. P., & Michahelles, F. (2011, September). Understanding social media marketing: a case study on topics, categories and sentiment on a Facebook brand page. In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments* (pp. 175-182). ACM.
4. Cvijikj, I. P., & Michahelles, F. (2013). Online engagement factors on Facebook brand pages. *Social Network Analysis and Mining*, 3(4), 843-861.
5. Cvijikj, I. P., Spiegler, E. D., & Michahelles, F. (2011, October). The effect of post type, category and posting day on user interaction level on Facebook. In *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third International Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on* (pp. 810-813). IEEE.
6. Dasgupta, S. S., Natarajan, S., Kaipa, K. K., Bhattacharjee, S. K., & Viswanathan, A. (2015, October). Sentiment analysis of Facebook data using Hadoop based open source technologies. In *Data Science and Advanced Analytics (DSAA), 2015. 36678 2015. IEEE International Conference on* (pp. 1-3). IEEE.
7. De Valck, K., Van Bruggen, G. H., & Wierenga, B. (2009). Virtual communities: A marketing perspective. *Decision Support Systems*, 47(3), 185-203.

8. De Vries, L., Gensler, S., & Leeﬂang, P. S. (2012). Popularity of brand posts on brand fan pages: An investigation of the effects of social media marketing. *Journal of Interactive Marketing*, 26(2), 83-91.
9. Del Giudice, M., Della Peruta, M. R., & Carayannis, E. G. (2013). *Social media and emerging economies: Technological, cultural and economic implications*. Springer Science & Business Media.
10. Dellarocas, C. (2006). Strategic manipulation of internet opinion forums: Implications for consumers and firms. *Management science*, 52(10), 1577-1593.
11. Goh, K. Y., Heng, C. S., & Lin, Z. (2013). Social media brand community and consumer behavior: Quantifying the relative impact of user-and marketer-generated content. *Information Systems Research*, 24(1), 88-107.
12. Habernal, I., Ptáček, T., & Steinberger, J. (2014). Supervised sentiment analysis in Czech social media. *Information Processing & Management*, 50(5), 693-707.
13. Hamouda, S. B., & Akaichi, J. (2013). Social networks' text mining for sentiment classification: The case of Facebook's statuses updates in the 'Arabic Spring' era. *International Journal Application or Innovation in Engineering and Management*, 2(5), 470-478.
14. Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1993). Emotional contagion: Studies in emotion and social interaction.
15. Heinonen, K. (2011). Consumer activity in social media: Managerial approaches to consumers' social media behavior. *Journal of Consumer Behaviour*, 10(6), 356-364.
16. Hilbe, J. M. (2011). *Negative binomial regression*. Cambridge University Press.

17. Hollebeek, L. D. (2011). Demystifying customer brand engagement: Exploring the loyalty nexus. *Journal of marketing management*, 27(7-8), 785-807.
18. Jurek, A., Mulvenna, M. D., & Bi, Y. (2015). Improved lexicon-based sentiment analysis for social media analytics. *Security Informatics*, 4(1), 9.
19. Kang, H., Yoo, S. J., & Han, D. (2012). Senti-lexicon and improved Naïve Bayes algorithms for sentiment analysis of restaurant reviews. *Expert Systems with Applications*, 39(5), 6000-6010.
20. Khobzi, H., & Teimourpour, B. (2014). How significant are users' opinions in social media?. *International Journal of Accounting & Information Management*, 22(4), 254-272.
21. Luarn, P., Lin, Y. F., & Chiu, Y. P. (2015). Influence of Facebook brand-page posts on online engagement. *Online Information Review*, 39(4), 505-519.
22. Mangold, W. G., & Faulds, D. J. (2009). Social media: The new hybrid element of the promotion mix. *Business horizons*, 52(4), 357-365.
23. Meire, M., Ballings, M., & Van den Poel, D. (2016). The added value of auxiliary data in sentiment analysis of Facebook posts. *Decision Support Systems*, 89, 98-112.
24. Moore, D. S., & McCabe, G. P. (1989). *Introduction to the Practice of Statistics*. WH Freeman/Times Books/Henry Holt & Co.
25. Moro, S., Rita, P., & Vala, B. (2016). Predicting social media performance metrics and evaluation of the impact on brand building: A data mining approach. *Journal of Business Research*, 69(9), 3341-3351.

26. Muntinga, D. G., Moorman, M., & Smit, E. G. (2011). Introducing COBRAs: Exploring motivations for brand-related social media use. *International Journal of advertising*, 30(1), 13-46.
27. Neri, F., Aliprandi, C., Capeci, F., Cuadros, M., & By, T. (2012, August). Sentiment analysis on social media. In *Proceedings of the 2012 International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2012)* (pp. 919-926). IEEE Computer Society.
28. Ngoc, P. T., & Yoo, M. (2014, February). The lexicon-based sentiment analysis for fan page ranking in Facebook. In *Information Networking (ICOIN), 2014 International Conference on* (pp. 444-448). IEEE.
29. Ortigosa, A., Martín, J. M., & Carro, R. M. (2014). Sentiment analysis in Facebook and its application to e-learning. *Computers in Human Behavior*, 31, 527-541.
30. Pham, T. N., Vuong, T. H., Thai, T. H., Tran, M. V., & Ha, Q. T. (2016). Sentiment Analysis and User Similarity for Social Recommender System: An Experimental Study. In *Information Science and Applications (ICISA) 2016* (pp. 1147-1156). Springer Singapore.
31. Rieder, B. (2013). Studying Facebook via data extraction: the Netvizz application. In *Proceedings of the 5th Annual ACM Web Science Conference* (pp. 346-355). ACM.
32. Sabate, F., Berbegal-Mirabent, J., Cañabate, A., & Lebherz, P. R. (2014). Factors influencing popularity of branded content in Facebook fan pages. *European Management Journal*, 32(6), 1001-1011.
33. Salehan, M., & Kim, D. J. (2016). Predicting the performance of online consumer reviews: A sentiment mining approach to big data analytics. *Decision Support Systems*, 81, 30-40.

34. Shen, B., & Bissell, K. (2013). Social media, social me: A content analysis of beauty companies' use of facebook in marketing and branding. *Journal of Promotion Management*, 19(5), 629-651.
35. Steuer, J. (1992). Defining virtual reality: Dimensions determining telepresence. *Journal of communication*, 42(4), 73-93.
36. Tafesse, W. (2015). Content strategies and audience response on Facebook brand pages. *Marketing Intelligence & Planning*, 33(6), 927-943.
37. Tang, T., Fang, E., & Wang, F. (2014). Is neutral really neutral? The effects of neutral user-generated content on product sales. *Journal of Marketing*, 78(4), 41-58.
38. Thelwall, M. (2013). Heart and soul: Sentiment strength detection in the social web with sentistrength. *Proceedings of the CyberEmotions*, 1-14.
39. Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., & Kappas, A. (2010). Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, 61(12), 2544-2558.
40. Thongmak, M. (2015, June). Engaging Facebook users in brand pages: Different posts of marketing-mix information. In *International Conference on Business Information Systems* (pp. 299-308). Springer International Publishing.
41. Troussas, C., Virvou, M., Espinosa, K. J., Llaguno, K., & Caro, J. (2013, July). Sentiment analysis of Facebook statuses using Naive Bayes classifier for language learning. In *Information, Intelligence, Systems and Applications (IISA), 2013 Fourth International Conference on* (pp. 1-6). IEEE.

42. Turney, P. (2002). Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the Association for Computational Linguistics (ACL)* (pp. 417–424)
43. Van Dijk, J. A. (2006). Digital divide research, achievements and shortcomings. *Poetics*, 34(4-5), 221-235.
44. Weichselbraun, A., Gindl, S., & Scharl, A. (2011, October). Using games with a purpose and bootstrapping to create domain-specific sentiment lexicons. In *Proceedings of the 20th ACM international conference on Information and knowledge management* (pp. 1053-1060). ACM.
45. Wood, C. A., Ray, S., & Messinger, P. (2013, August). Leaving the tier: an examination of asymmetry in pricing patterns in online high tech shops. In *International Conference on Electronic Commerce* (pp. 63-73). Springer Berlin Heidelberg.
46. Yang, X., Yang, G., & Wu, J. (2016). Integrating rich and heterogeneous information to design a ranking system for multiple products. *Decision Support Systems*, 84, 117-133.
47. Zhang, K., Bhattacharyya, S., & Ram, S. (2014, September). Empirical analysis of implicit brand networks on social media. In *Proceedings of the 25th ACM conference on Hypertext and social media* (pp. 190-199). ACM.
48. Zhang, K., Bhattacharyya, S., & Ram, S. (2016). Large-Scale Network Analysis for Online Social Brand Advertising. *MIS Quarterly*, 40(4), 849-868.