International Journal of Management, Accounting and Economics
Vol. 8, No. 9, September, 2021
ISSN 2383-2126 (Online)

www.ijmae.com

Case Study

Customers' Emotions towards Digikala Online Retail Services during COVID Pandemic: A Twitter Dataset Based Analysis

Nasim Bagheripour Salamat and Maryam Pournasir Rudbaneh[1]
Department of Mathematics, Lahijan Branch, Islamic Azad University, Lahijan, Iran

Abstract

The main purpose of this research is to identify the concerns of Digikala's online customers during Corona Crisis, which has been discussed and shared among Twitter users. This study collects hashtag-based tweets related to digital goods that cover the 2020 period. To analyze the tweets, text analytical methods such as thematic modeling, emotion analysis, and word clouds are used. In this way, the most important topics discussed on Twitter have been discovered, then customer dissatisfaction has been determined. Afterward, the obtained results were compared with previous in the field of online retail services. Further, noticed that the trend of unforeseen trending hashtags related to digital goods in the last year, which identifies other factors that users are discussing. This study gives us an insight into how the brand operates in the Covid-19 crisis. Analysis of this information can be used to improve online retail services and access to customer needs.

Keywords: Covid 19, Emotion Analysis, Digikala, Twitter, Online Retail

[1] *Corresponding author's Email: maryampournasirrudbaneh@gmail.com*

Introduction

Over the last few years, cyberspace and virtual networks have evolved significantly, with many people using these spaces and social networks for a variety of purposes and constantly sharing a large amount of this information in various forms, including image, video, and text(Kim, Breslin, Chao, & Shu, 2011). Given this large amount of data generating daily, new areas such as data processing and storage were emerged, necessitating studies in the field of data mining and text mining. These techniques are in fact the process of retrieving knowledge from data that is a very pivotal task. Therefore, we require appropriate techniques to operate and analyze these data.

Twitter is one of the most widely used social media today, and since a significant portion of the content generated in the social media Twitter is text data, the analysis of these texts can give a very good perspective of individuals' ideas and views on various topics to researchers and business owners(Salloum, Al-Emran, Monem, & Shaalan, 2017). A remarkable behavior observed among many customers is that they are interested in sharing their satisfaction or dissatisfaction with a service or product in a larger space in addition to relatively traditional methods such as contacting customer service. As a result, their messages and comments are seen by a lot of people in a very short time and they receive a flood of similar and opposite messages and comments, which are considered as both challenge and opportunity for businesses and enable retailers or businesses to use this data to gain meaningful knowledge.

On the other hand, the novel coronavirus (COVID-19) pandemic with a profound impact on people's lifestyles, including adherence to new rules such as quarantine, traffic restrictions and slogans has led many people to turn to online shopping, and the recent huge volume of new customers has simultaneously created many opportunities for online retailers as well as many challenges, as they also involve the same norms and restrictions. In this project, we evaluate the ideas and comments of Digikala online retail customers during this crisis.

The coronavirus pandemic has imposed an emergency situation on people all over the world and has challenged online retailers, shutting down many businesses and downsizing many of their workforces, and even disrupting trade and industry. Quarantine has also led, for example, people to enter the online shopping space as consumers who have never considered it a way to purchase.

Twitter is an online social media platform that has active users worldwide who daily share a lot of content with others(Java, Song, Finin, & Tseng, 2007). In this study, we used the Latent Dirichlet Allocation (LDA) thematic modeling approach(Xue et al., 2020), which is an unsupervised learning method, to identify latent patterns in emotions that are indexed through network analysis and word cloud. Finally, the results were compared with the results obtained from previous literature on online retail services.

Emotion and opinion analysis helps companies, organizations, and businesses find out what users think about them, as well as informing customers of what customers think of the product or service offered. In fact, in order to attract, retain customers, it is necessary to value their opinions and needs, and try to eliminate them and obtain their satisfaction by improving the quality of products and services. Analyzing users' opinions and feelings on Twitter can go a long way in helping customers understand their businesses and enhance their market share and competitive advantage.

This study mainly aims to analyze the feelings of consumers based on their ideas about Digikala online retail during the COVID-19 crisis and to explore the important topics that Twitter users often talk about. This study also attempts to answer the question: What are the most commonly discussed and shared topics among Twitter users about online retail brand?

Literature Review

Online reviews can be used as a major and pivotal information source for researchers to better understand consumers' needs and demands (Chau & Xu, 2012). Their analysis can also facilitate the quality of products and services by identifying customers' demands and new marketing strategies (Loureiro & Kastenholz, 2011). Reviews allow other individuals to put aside a vague products or services and directly benefit from the experiences of other consumers. Furthermore, some companies encourage customers to leave online reviews for products or services by offering discounts and services (Guo, Barnes, & Jia, 2017).

Social media and technological concepts have revolutionized business-to-business (B2B), business-to-consumer (B2C), and consumer-to-consumer (C2C) communications (Kietzmann, Hermkens, McCarthy, & Silvestre, 2011). The Internet has evolved from a one-way information dissemination platform to a participatory platform, enabling people to be self-media (Li & Wang, 2011). For instance, travelers widely for travel planning utilize search, sharing experiences and travel stories in blogs and online communities, etc. (Leung, Law, van Hoof, & Buhalis, 2013).

In addition, online customer reviews have become a significant tool for companies and customers to receive and provide feedback on products and services. Marketers utilize online customer surveys as a means of driving firm's and also evaluate customers' feelings towards the brand and their level of real satisfaction (Willemsen, Neijens, Bronner, & Ridder, 2011). Valuable information and knowledge derived from social media are useful for business intelligence at different levels of management in organizations and play a pivotal role in business success (Ramanathan, Subramanian, & Parrott, 2017) and branding (Cawsey & Rowley, 2016) .For example, (Cawsey & Rowley, 2016) argued that social media is a strategic approach to branding and emphasized that brand image and increasing brand awareness are the most common motivations for presence in social media. (Wang, Yu, & Wei, 2012) suggested that customers use social media to socialize and talk about products, and found that a clear relationship has a positive relationship with customers' insights about a brand's products. Researchers in (Salkhordeh, 2011) provided two results of using social media as follows:

1. A valuable tool for collecting customer feedback to persuade new and existing customers that can build an effective and powerful relationship with the customer to increase brand loyalty.

International Journal of Management, Accounting and Economics
Vol. 8, No. 9, September, 2021
ISSN 2383-2126 (Online)

www.ijmae.com

2. Improper application reduces customer trust and lowers the brand value.

Online social media refers to social activities occurring in the online environment, in which Internet users communicate with each other online. The emergence of these virtual communities provides a platform for online consumers to share and exchange their perspectives and information on products or services. This provides great opportunities for retailers, especially online businesses. Moreover, online consumers develop online trust through interactions in virtual communities (Balakrishnan, Dahnil, & Yi, 2014) .

Datasets and Pre-Processing

In this article, data are sourced from (Phantombuster, 2021) and have collected from Twitter. Thus, related hashtag-based tweets were collected in two rounds of three and six months of 2020, of which 70% are related to the six-month period and 30% are related to the next three months. The initial data for the hashtags Digikala and Digi-Kala with a 1 million tweets were received in csv format.

Pre-processing is one of the most important steps when working with data. The better the data is prepared in the pre-processing stage, the better the results will be. In fact, tweets at the time of collection include unnecessary words such as prepositions, numbers, photos, emojis, pronouns, links, etc. that need to be removed from tweets. In addition, we have many retweets in the dataset that must also be deleted. The initial pre-processing of tweets is done in Excel software. One of the branches of preprocessing is refining, which is done by removing the extra rows and columns by filtering the main matrix.

In the first stage, the reduction of the column dimension is done. It deals with removing the extra columns of Figure (1) like C and D but remaining column B. We also remove columns related to the names and information of the tweeters because they are unnecessary information. Also, 25% of the data was related to the digital goods support department, which are also deleted and related to the step of reducing the rows (Figure 2).

	A	B	C	D	E	F
750	Sat Jan 09 15:17:22 +0000 202	@Zoohreeh نکردم ،یکم با استفاده اش مشکل دارم	https://twitter.com/nilgoneabi	https://twitter.com/nilgo	2021-01-13T22:33:4	دیجی_کالا
751	Thu Jan 07 19:47:39 +0000 202	@Saman_830 الا نباشه جنس ها تو از کجا میگیری	https://twitter.com/Nima74995	https://twitter.com/Nima	2021-01-13T22:33:4	دیجی_کالا
752	Wed Jan 06 20:59:12 +0000 20	هر کارفرمای ایرانی --->gt;& ی پروژه میخوام مثل دی	https://twitter.com/NimaRastg	https://twitter.com/Nima	2021-01-13T22:33:4	دیجی_کالا
753	Sun Jan 10 23:10:07 +0000 202	@minoodarya @SSuereh زینه خود کالا بکشیم؟	https://twitter.com/nimche_ra	https://twitter.com/nimc	2021-01-13T22:33:4	دیجی_کالا
754	Wed Jan 13 19:33:48 +0000 20	شاید باورتون نشه ولي يكي از تفريحات من اینه که می	https://twitter.com/nimetamar	https://twitter.com/nime	2021-01-13T22:33:4	دیجی_کالا
755	Sun Jan 10 10:34:55 +0000 202	@salam70519015 اجناس بد و گرون هم زیاد داره	https://twitter.com/nimetamar	https://twitter.com/nime	2021-01-13T22:33:4	دیجی_کالا
756	Wed Jan 13 10:41:57 +0000 20	هرسری میرم دیجی کالا قسمت لپتاپو باز میکنم میبین	https://twitter.com/nImKHa3	https://twitter.com/nImK	2021-01-13T22:33:4	دیجی_کالا
757	Thu Jan 07 05:49:16 +0000 202	هنوز از دیجی کالا خرید نکردم چون نمیدونم مالک و	https://twitter.com/nkm00710	https://twitter.com/nkm0	2021-01-13T22:33:4	دیجی_کالا
758	Wed Jan 13 20:10:13 +0000 20	@aminoroaya313 مین دوره ۴۰ ساله جون گرفتن	https://twitter.com/nkm00710	https://twitter.com/nkm0	2021-01-13T22:33:4	دیجی_کالا
759	Sun Jan 10 08:29:27 +0000 202	فقط من چشمم باکس دیجی کالا رو گرفت؟ تا باکستم ک	https://twitter.com/Noghteh_	https://twitter.com/Nogh	2021-01-13T22:33:4	دیجی_کالا
760	Fri Jan 08 08:00:41 +0000 2021	@bazhool @Macmaneman1 یگم زده۲۵ تومن	https://twitter.com/Nonecolou	https://twitter.com/None	2021-01-13T22:33:4	دیجی_کالا
761	Sun Jan 10 07:51:12 +0000 202	فاز فروشنده های لباس ، کفش تو دیجی کالا نمیفهمم آ	https://twitter.com/omid_sh1	https://twitter.com/omid	2021-01-13T22:33:4	دیجی_کالا
762	Tue Jan 12 12:46:05 +0000 202	@SiavashGhambari یجی کالا قیمت لپ تاپ ببین	https://twitter.com/omidziadz	https://twitter.com/omid	2021-01-13T22:33:4	دیجی_کالا

Figure 1. Reducing column dimensions

International Journal of Management, Accounting and Economics
Vol. 8, No. 9, September, 2021
ISSN 2383-2126 (Online)

www.ijmae.com

	A	B	C
242	Tue Jan 12	@nazaniin_am 😊.امیدواریم لبخندی که زمان رسیدن سفارشتون از دیجی‌کالا می‌زنین، سختی انتظار رو تلافی کنهسلام نازنین جان	https://twitter.com/digikalasupport
243	Thu Jan 07	@selenophiiille کسفارش رو در دایرکت ارسال کنین تا بتونیم وضعیت سفارش رو بررسی کنیم و نتیجه رو بهتون اطلاع بدیمسلام آیدا جان	https://twitter.com/digikalasupport
244	Sun Jan 10	@Hamik91 برای بررسی و ارائه توضیح دقیق‌تر لینک کالا رو به ما بدین.	https://twitter.com/digikalasupport
245	Sun Jan 10	@homayounmg مبارکتون باشه و به‌خوشی استفاده کنین	https://twitter.com/digikalasupport
246	Wed Jan 1	@Lostisolated می‌کنیم. لطفا شماره سفارش‌تون رو دایرکت برای ما ارسال کنین تا این موضوع رو دقیق‌تر بررسی و پیگیری کنیمسلام دوباره	https://twitter.com/digikalasupport
247	Mon Jan 1	@frz_kmn_70 و مشتری رو تامین کنه. کالاهایی که برچسب قیمت دارن و قیمت مصرف‌کننده مشخصه امکان فروش با قیمت بیشتر ندارنسلام	https://twitter.com/digikalasupport
248	Thu Jan 07	@stabillox چالش‌ها ظرفیت ارسال ما رو کاهش داده و باعث شده بازه‌های ارسال کمی طولانی‌تر از شرایط عادی باشن.	https://twitter.com/digikalasupport
249	Mon Jan 1	@frz_kmn_70 اگه اطلاعاتی دارین دایرکت برامون ارسال کنین تا پیگیری کنیم. در صورت تایید با فروشنده برخورد و قیمت کالا اصلاح میشه	https://twitter.com/digikalasupport
250	Sun Jan 10	@imadumbghost ☺ میزبانی از شما باعث افتخارمونه آرمین جان	https://twitter.com/digikalasupport
251	Mon Jan 1	@abiyanaki امیدواریم در خریدهای بعدی بتونیم تجربه بهتری رو براتون رقم بزنیم.	https://twitter.com/digikalasupport
252	Fri Jan 08	@esi_the_20 لطف می‌کنین لینک نمونه کالایی که داره با قیمت بالاتر فروخته می‌شه، در دایرکت ارسال کنین؟	https://twitter.com/digikalasupport
253	Mon Jan 1	@SphrDiary خریدهاشون رو اینترنتی انجام میدن، در این شرایط ما هم در تلاشیم تا همه سفارش‌ها رو به موقع برای مشتری‌ها ارسال کنیمسلام	https://twitter.com/digikalasupport
254	Thu Jan 07	@stabillox بار فروشنده به انبار دیجی‌کالا، شفت‌بندی و رفت‌وآمد همکاران تیم عملیات و همچنین تردد ماموران ارسال ما تحت تاثیر قرار گرفته	https://twitter.com/digikalasupport
255	Sun Jan 10	@alirezapt برای وصل شدن به پشتیبانان تلفنی کافیه کمی صبوری کنینسلام علیرضا جان (:	https://twitter.com/digikalasupport
256	Fri Jan 08	@jigsaw89295108 @Digikalacom نده‌هاست تا تجربه خرید و فروش رو میسر کنیماین موارد رو به همکاران پشتیبانی ما اطلاع دادین؟سلام	https://twitter.com/digikalasupport
257	Tue Jan 12	@shabgard_98 در دایرکت با ما در ارتباط باشین. اگه سوالی هست در خدمتیمسلام	https://twitter.com/digikalasupport
258	Thu Jan 07	@Passionfruitam لطفا به برادرتون سلام ما رو برسونین و بگین باعث افتخارمونه که روی ما حساب می‌کنه	https://twitter.com/digikalasupport
259	Sun Jan 10	@miladjef ها رو به همراه یک کد تخفیف عذرخواهی بازگشت می‌دیم.برای عدم تامین به موقع کتاب ازتون عذرخواهی می‌کنیمسلام میلاد جان	https://twitter.com/digikalasupport
260	Sun Jan 10	@miladjef امیدواریم بتونیم این میزان عدم تامین رو هم به حداقل برسونیم و تجربه خوشایندی از خرید برای مشتری‌ها رقم بزنیم.	https://twitter.com/digikalasupport

Figure 2. Reducing raw dimensions

The next step is to delete spams. Spams are those tweets which do not contain specific information and can be deleted. We also remove and delete tweets that contain only attachments such as a photo or video without text content with the hashtag Digikala. Some of the tweets contained inappropriate words that we put in the so-called insult subcategory and are not counted as spams.

Thematic Modeling

One of the popular tools for extracting topics on social networks and facilitating researchers' understanding of large unstructured datasets is thematic modeling. To find the topics, we use the LDA (Linear discriminant analysis) algorithm, a statistical model used to identify the main topics discussed in a set of documents. This model receives a set of documents and cluster a set of topics. Each topic contains a set of words. This requires defining the number of topics represented by k for modeling distributions.

Latent Dirichlet Allocation

The LDA is a potentially productive model of a data set. First introduced in 2003 (Blei, Ng, & Jordan, 2003) ،it is one of the most popular methods in subject modeling. The LDA indicates topics with word probabilities. The words that are most likely in each topic usually give us a good idea of what the topic is about. The LDA is an unsupervised method for modeling large datasets that is the most common method of subject modeling. The LDA assumes that each document can be shown as a potential distribution on hidden subjects(Chen, Zou, Cheng, & Xie, 2020), and that distribution distributes the subject in all documents of a former common Dirichlet. Each topic in the LDA model is also shown as a possible distribution on words, and the word distribution of topics also shares a previous common Dirichlet.

Having a set D consisting of M document, with document d containing N_d words $(d\in 1,\ldots M)$، D models can be made according to the following generating process: A) Choose a polynomial distribution of ϕ_t for the subject $t(t\in\{1, \ldots, T\})$ from the Dirichlet distribution with the parameter β.

B) Choose a polynomial distribution θ_d for the document d ($d \in\{1, \ldots, M\}$) from the Dirichlet distribution or parameter α.

C) for a word w_n ($n \in\{1,\ldots,N_d\}$) in document d,

1. Select a z_n subject from θ_d .

2. Choose a word w_n from ϕ_{zn}.

In the above generative process, words in documents are the only observed variables while the rest are hidden variables (θ and ϕ) and hyper parameters (α and β). In order to deduce the hidden variables and hyper parameters, the probability of the observed data D is calculated as follows and maximized:

$$P(D|\alpha.\beta) = \prod_{d=1}^{M} \int p(\theta_d|\alpha)(\prod_{n=1}^{N_d} \sum_{z_{dn}} p(Z_{dn}|\theta_d) P(w_{dn}|z_{dn}.\beta)) d\theta_d \tag{1}$$

The specified parameters are α for the previous Dirichlet topic and the word distribution on the topics, which is called the Dirichlet β distribution. T is the number of topics, M is the document number and N is the word size. Dirichlet polynomial pairs are considered as (α, β) for thematic distributions at the data set level. The pair of Dirichlet - polynomials for subject-word distributions are called (β, ϕ). Θ_d variables are document related variables that are sampled in each document. The z_{dn} w_{dn} variables are word related variables and are sampled for each word in each text document(Vayansky & Kumar, 2020).

When importing documents (tweets) from Python after pre-processing to apply the LDA, the data will be processed into tables of pre-processed text and there will be a record. Now each word in each record needs to be semantic. In this way, the words of each record are automatically checked and matched using the libraries defined for the Persian language that are used and categorized in Python. Some words are adverbial, complementary, pronoun and object, so they are not used to make meta-words, so they are removed and only nouns and verbs with emotional meanings remain. To facilitate work in the Python environment, many words were manually rooted in the preprocessing stage so there is no re-rooting. Then some of the words match the words of one of the dictionaries and some of them match the words of another dictionary until all the words in the document are matched. In this way, the words that match a dictionary form a cluster. Now, if a word in a document is repeated many times, it has a higher frequency, so in a word cloud, which is based on frequency, it will be in the center of the cloud, indicating the importance of the word in the document. If a word is in the center of the cluster, it is based on the frequency of the word in the document. All code used in the Python environment uses packages in Python that can be called by calling any of the packages. Packages used in Python are Tweepy, Pandaz, Sklearn, WordCloudFa, Matplotlib and, pyLDAvis.

International Journal of Management, Accounting and Economics
Vol. 8, No. 9, September, 2021
ISSN 2383-2126 (Online)

www.ijmae.com

Analysis and Results

In this section, the results of the LDA thematic model analysis from their tweet collections are presented. it has been demonstrated the results using word cloud on key topics and related topics to identify important factors for online retail. we compared the factors found in the LDA analysis with existing online retail literature. In this study, we selected the total number of topics for clustering from 2 to 5 topics, which was finally enough to present the display of K =2 subject, but the result of clustering performance up to k = 5 will be shown in the next section based on f1 criteria. In this section, we will deal with the results of clusters and word clouds. Figures 3 and 4 are the clustering results of the last three months. Figure 3 is related to the first cluster (first topic) and Figure 4 is related to the second topic. As can be seen, in these two forms, the center of the clusters, which are the largest words in the word cloud, are the words price and insult, respectively. But first, before the secondary studies, it is important to mention a few points.

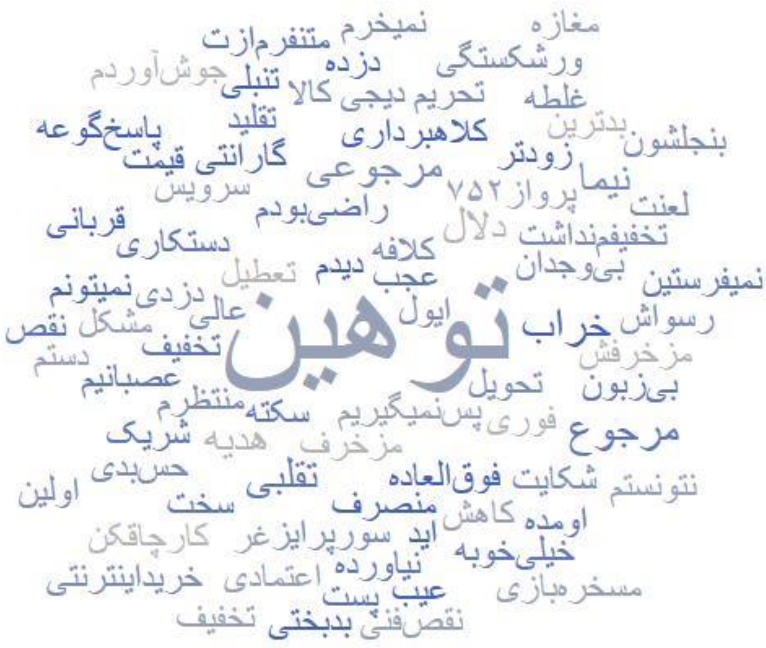


Figure 3. Result of the first cluster clustering in the data of the last three months

First, part of the dissatisfaction was due to the sudden devaluation of the national currency and the subsequent rise in prices of goods and products, and second, the dissatisfaction of public opinion with the events and sexual harassment, due to one of the managers of Digikala, which was widely discussed on Twitter. It damaged the brand value and made users distrustful. The third case was the start of a new phase of quarantine and restrictions, which in turn led to a significant increase in online shopping requests, and it is understandable that dissatisfaction will increase because the network was not prepared to deal with this volume of customers and causes a decline. Quality in services, for example, complaints Figure 3 like errors in sending and increasing delivery delay are often seen in tweets. According to the word clouds, we realize that the most important areas of dissatisfaction in the last three months are related to prices, which, as we mentioned, can be an important factor in the changes of the national currency. But as an

International Journal of Management, Accounting and Economics
Vol. 8, No. 9, September, 2021
ISSN 2383-2126 (Online)

www.ijmae.com

influential factor along with other words related to retail services according to previous studies are listed in the following tables.

Figure 4. Result of the second cluster clustering in the data of the last three months

Table 1. Clustering results of the first cluster of data of the last three months

Subject	Word	Parameter	Description	Source
Product	Price	Product price	Product selling price	(Fang et al., 2014)
	unavailable	Product availability	Availability of products in the store for potential customers	(Chang et al., 2004)
	used	Product experience	Product comments and evaluations	(Brakus et al., 2009)
Delivery services	mistake	Making mistake and delay in sending	Making mistake and delay in sending product	(Xu, Wang, Li, & Haghighi, 2017)
	Taking too long	Delivery schedule	Timely delivery, Receive the product on time	(Cheung et al., 2008; Luo et al., 2012)

According to Table 1 and the word cloud of Figure 1, it is clear that words such as" non-available" and "used" are related to dissatisfaction of the product section and the "mistake" and "taking too long" are related to the delivery section. But the center of the cluster, which is "price", has the most weight in the cluster, does not indicate a negative or positive feeling, but it was the most discussed issue at the time, which, as we mentioned, seems normal given the reducing the value of currency rate. Words such as "dissatisfaction", "distrust" and "insult" indicate general dissatisfaction and words such as "excellent", "discount", and "excellent" have positive meanings. But words like "addict" were actually related to a few tweets that indicate that customers are satisfied with your purchase from Digikala, and that you are addicted to online shopping, and it does not necessarily have a negative connotation here. Even words like "price", although emotionally neutral in meaning, are included in this cluster because they are placed next to other words that have a positive or negative meaning. Although Digikala has seen an increase in the volume of purchases by users during this period, for various reasons, most of the words that have more weight in the cluster have negative meanings.

According to Figure 2 and Table 2, we see that in the second cluster head, the word "insult" has a very high weight because in preprocessing we converted about 30 insulting words into "insult". Other words like "well done", "Excellent", "Very Good", "Surprise", and "Gift Discount" are positive, and most other words are negative. However, in this cluster, according to the table of words such as "counterfeit", "broken", and "defective" are related to the product area and words such as "waiting", do not send and earlier in the field of delivery of goods and customer service section includes words such as responsive, reference, it is a guarantee. Since words like "hate", "bad feeling", and "clutter" are next to "insult" in this cluster, these words are most likely selected from a dictionary and require clusters so that the cluster does not focus so much on negative words and emotions.

The clustering results for the last six months are shown in Figures 3 and 4. Figure 3, as shown in Figure 2, shows that in this time period, the word "insult" has a lot of weight, which was mentioned above with the same logic. In the second place is "omitted" and "earlier", and some other words have almost the same frequency. A noteworthy figure in this figure refers to a time when users on Twitter were talking about replacing Digikala with a platform with similar services. Words like "troob", "hojrak", "marketing", and "raghib govah" are competitors.

Figure 4 shows the result of the second clustering in the last six months, which again, like Figure 2, deals with the issue of high weight "cost". Words and topics related to economic issues are mostly seen in this form. By examining these forms, it shows that clustering is meaningful and appropriately segregates related topics and places them in a cluster, not just on the basis of frequency. But due to the small number of topics and limiting it to two topics, unrelated words can also be seen in this form.

Table 2. Results of the second clustering data of the last three months

Subject	Word	Parameter	Description	Source
Product	Cost	Product price	Product selling price	(Fang et al., 2014) (Luo et al., 2012)
	Fake	Product knowledge	Knowledge about the product	)Lim & Chung, 2011)
	Broken	Product experience	Consumers directly or indirectly express their opinions and evaluate products	(Brakus, Schmitt, & Zarantonello, 2009)
	Defect / technical defect			
	Defect			
Delivery services	Waiting	Delivery schedule	Timely delivery, receiving the shipment according to the guaranteed time	Cheung, Lee, & Rabjohn, 2008)
	Not Sent		Timely delivery, receiving the shipment according to the guaranteed time	(Luo et al., 2012)
	Sooner			
	Immediate			
	Post			
Customer Services	Responsive	responsiveness	Immediate response to customer requests and complaints	(Francis & White, 2002)
	Reference	Return and Refund	Simplicity in return the product or refunding it	(Ahn et al., 2004) (King et al., 2014)
	Warranty			
	Service	Customer support	Following up the service after receiving the cost of received feedbacks from the staff	(Francis, 2007) (J. E. Francis, 2009)
	Complaint	Empathy	Understand customer requirements and references	(Ahn et al., 2004)

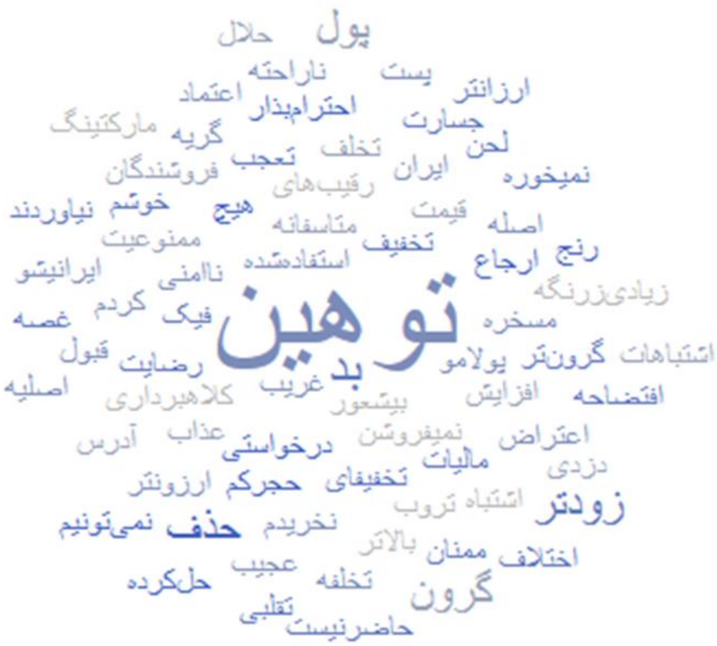
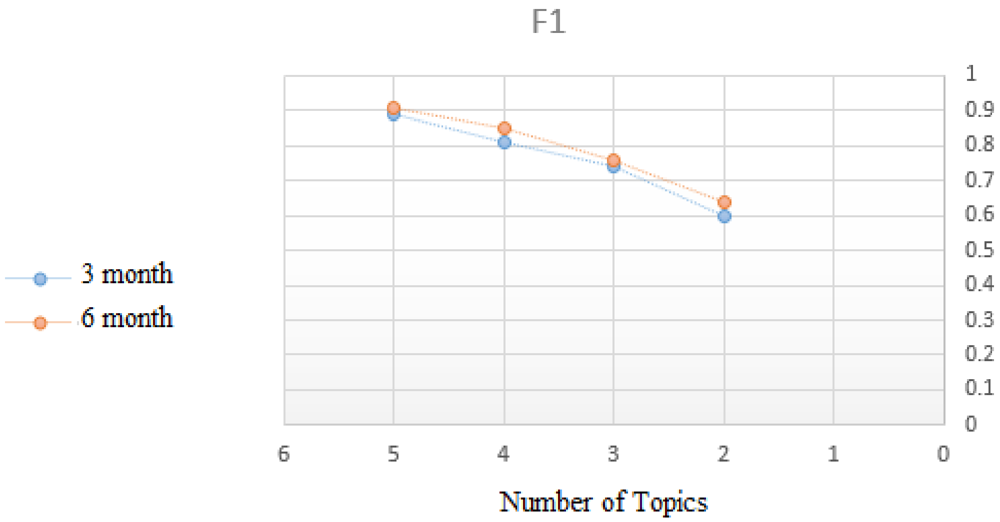


Figure 5. The result of the first cluster in the last six months

Figure 6. The result of the second cluster in the last six months



Polt 1. Comparison of F1 results of data clustering values over time

The results of clustering clusters in online retail services are as follows:

- Clustering of first clusters of the first three month shows that user's main concern area was about "product" and "delivery related services".

- Clustering results of the second clusters of the first three months shows that concern area in this cluster is mainly about "product", "delivery related services" with more components than the first clusters. and many words related to "customer service" was obtained.

- In the clustering of the first clusters of the first six months' increase of words related to "product" and after that "delivery related services" and "customer service" were obtained.

- Clustering results of the second clusters of the second six months showed that "product", "delivery related services" and "customer service" were the main concern areas and were more discussed among twitter users.

With a closer look to the issue that was just discovered, one can receive important managerial topics: to pay more attention to customer's needs for creating new strategies to overcome those needs and to create new ways for development. Creating a bond with customer and pay more attention to customer relationship has to be a priority for firms. Social media platforms have to be utilized for achieving the desired results. Social media platforms are places where customers talk about their needs and concern areas, therefor effective social media management allows companies to listen more and pay more attention to these needs. Unlike traditional approaches where firms tend to talk more and promote themselves rather than listening. Social media era gave this opportunity to businesses to talk less and listen more which has more benefits to it.

Conclusion

The main goal of this study was to explore repeated topics and to identify customer's sentiments about Digikala online retailing by collecting fair data from Twitter. The reason why Twitter was chosen for this study is the presence of real and potential customers of online shopping according to age range. Also following up tweet threads are much simpler than other social media platforms like Instagram. To understand what topics are discussed often in this platform, a text mining clustering algorithm -which also considers term weights- called LDA was used to classify clusters based on their topics. To process tweets, two different time ranges in 2020 were used. The year which was started with Covid-19 in Iran and faced businesses with several crises. But the data used for this study were collected after three months since the crisis started. Because we wanted the results to be fair, therefor, a three-month gap was considered for the retail to explore the dimensions of the accident. To evaluate the clustering performance of the two cluster clustering, word clouds were made. Center word of these clusters which had the most weights in the last three and six months was as followed:

- For three months: Price, insults, expensive, and face masks are related to online retail services.

- For six months: Unavailable, taking too long, wrong, product, refund, broken, original, responsiveness, guaranty, trust, etc.

Then F1 Evaluation criteria were used to evaluate the clusters themselves. Index results show that with five clusters, a 90 percent result can be reached. This means that the input data had at least five topics.

Referring to the F1 diagram, we find that the number of topics is at best 5 = k topics. Since we faced systemic limitations in this study, it is suggested that this number of topics be considered more than 5 categories. Also, the small number of topics and the fact that we only examined Digikala datasets, in future studies, the use of data from other companies, brands and competitors to compare areas of concern and discussion of users and customers will be very useful.

By paying attention to the topics that were discovered, important management issues can be obtained from the results and customers' needs can be considered more, and new strategies and development can be considered. Communicating with customers should be a priority for companies, and social media platforms should be fully utilized to achieve the desired result. Effective social media management gives companies access to a place where customers meet and talk, and enables them to understand and listen to their customers' concerns. Unlike traditional methods, in which companies tend to talk rather than listen, the age of social media provides companies with the opportunity to talk less and listen more.

References

Balakrishnan, B. K., Dahnil, M. I., & Yi, W. J. (2014). The impact of social media marketing medium toward purchase intention and brand loyalty among generation Y. Procedia-Social and Behavioral Sciences, 148, 177-185.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. the Journal of machine Learning research, 3, 993-1022.

Brakus, J. J., Schmitt, B. H., & Zarantonello, L. (2009). Brand experience: what is it? How is it measured? Does it affect loyalty? Journal of marketing, 73(3), 52-68.

Cawsey, T., & Rowley, J. (2016). Social media brand building strategies in B2B companies. Marketing Intelligence & Planning.

Chau, M., & Xu, J. (2012). Business intelligence in blogs: Understanding consumer interactions and communities. MIS quarterly, 1189-1216.

Chen, X., Zou, D., Cheng, G., & Xie, H. (2020). Detecting latent topics and trends in educational technologies over four decades using structural topic modeling: A retrospective of all volumes of Computers & Education. Computers & Education, 151, 103855.

Cheung, C. M., Lee, M. K., & Rabjohn, N. (2008). The impact of electronic word-of-mouth: The adoption of online opinions in online customer communities. Internet research.

Fang, Y., Qureshi, I., Sun, H., McCole, P., Ramsey, E., & Lim, K. H. (2014). Trust, satisfaction, and online repurchase intention. MIS quarterly, 38(2), 407-A409.

Francis, J. E. (2007). Internet retailing quality: one size does not fit all. Managing Service Quality: An International Journal.

Francis, J. E., & White, L. (2002). PIRQUAL: a scale for measuring customer expectations and perceptions of quality in internet retailing. K. Evans & L. Scheer (Eds.), 263-270.

Guo, Y., Barnes, S. J., & Jia, Q. (2017). Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation. Tourism management, 59, 467-483.

Java, A., Song, X., Finin, T., & Tseng, B. (2007). Why we twitter: understanding microblogging usage and communities. Paper presented at the Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis.

Kietzmann, J. H., Hermkens, K., McCarthy, I. P., & Silvestre, B. S. (2011). Social media? Get serious! Understanding the functional building blocks of social media. Business horizons, 54(3), 241-251.

Kim, H.-L., Breslin, J. G., Chao, H.-C., & Shu, L. (2011). Evolution of social networks based on tagging practices. IEEE Transactions on Services Computing, 6(2), 252-261.

Li, X., & Wang, Y. (2011). China in the eyes of western travelers as represented in travel blogs. Journal of Travel & Tourism Marketing, 28(7), 689-719.

Lim, B. C., & Chung, C. M. (2011). The impact of word-of-mouth communication on attribute evaluation. Journal of Business Research, 64(1), 18-23.

Loureiro, S. M. C., & Kastenholz, E. (2011). Corporate reputation, satisfaction, delight, and loyalty towards rural lodging units in Portugal. International Journal of Hospitality Management, 30(3), 575-583.

Phantombuster. (2021). Code-free automations and data extraction. Retrieved from https://phantombuster.com/

Ramanathan, U., Subramanian, N., & Parrott, G. (2017). Role of social media in retail network operations and marketing to enhance customer satisfaction. International Journal of Operations & Production Management.

Salkhordeh, P. (2011). Key issues in use of social networking in hospitality industry: 2009.

Salloum, S. A., Al-Emran, M., Monem, A. A., & Shaalan, K. (2017). A survey of text mining in social media: facebook and twitter perspectives. Adv. Sci. Technol. Eng. Syst. J, 2(1), 127-133.

Vayansky, I., & Kumar, S. A. (2020). A review of topic modeling methods. Information Systems, 94, 101582.

Wang, X., Yu, C., & Wei, Y. (2012). Social media peer communication and impacts on purchase intentions: A consumer socialization framework. Journal of interactive marketing, 26(4), 198-208.

Xu, X., Wang, X., Li, Y., & Haghighi, M. (2017). Business intelligence in online customer textual reviews: Understanding consumer perceptions and influential factors. International Journal of information management, 37(6), 673-683.

Xue, J., Chen, J., Chen, C., Zheng, C., Li, S., & Zhu, T. (2020). Public discourse and sentiment during the COVID 19 pandemic: Using Latent Dirichlet Allocation for topic modeling on Twitter. PloS one, 15(9), e0239441.

International Journal of Management, Accounting and Economics
Vol. 8, No. 9, September, 2021
ISSN 2383-2126 (Online)

www.ijmae.com

HOW TO CITE THIS ARTICLE

Bagheripour Salamat, N., Pournasir Roudbaneh, M. (2021). Customers' Emotions towards Digikala Online Retail Services during COVID Pandemic: A Twitter Dataset Based Analysis. *International Journal of Management, Accounting and Economics*, 8(9), 668-683.

DOI: 10.5281/zenodo.5854892

URL: https://www.ijmae.com/article_139544.html