Sentiment Analysis User Comments On E-commerce Online Sale Websites

Faezeh Forootan*
M.Sc. in Information Technology Engineering, Ghiaseddin Jamshid Kashani University, Abyek, Iran
Forootan_faezeh@yahoo.com

Mohammad Rabiei
Department of Electrical and Computer, Faculty of Computer, Eyvanekey University, Semnan, Iran
Mohammad.rabiei@uniud.it

Received: 2019/12/04 Revised: 2020/06/04 Accepted: 2020/07/06

***Abstract*— E-commerce websites, based on their structural ontology, provides access to a wide range of options and the ability to deal directly with manufacturers to receive cheaper products and services as well as receiving comments and ideas of the users on the provided products and services. This is a valuable source of information, which includes a large number of user reviews. It is difficult to check the bulk of the comments published manually and non-automatically. Hence, sentiment analysis is an automated and relatively new field of study, which extracts and analyzes people's attitudes and emotions from the context of the comments. The primary objective of this research is to analyze the content of users' comments on online sale e-commerce websites of handcraft products. Sentiment analysis techniques were used at sentence level and machine learning approach. First, the pre-processing steps and TF-IDF method were implemented on the comments text. Next, the comments text were classified into two groups of products and services comments using Support Vector Machine (SVM) algorithm with 99.2% accuracy. Finally, the sentiment of comments was classified into three groups of positive, negative and neutral using XGBoost algorithm. The results showed, 95.23% and 95.12% accuracies for classification of sentiments in comments about products and services, respectively.**

Keywords—Machine Learning; Opinion Mining; Online Reviews; Sentiment Classification; TF-IDF; Xgboost

1. INTRODUCTION

The growth of e-commerce and, consequently e-commerce websites, had a significant impact on people's life. In this situation, many people share their opinions about online products and services as well as services offered by e-commerce websites during the shopping process [1]. Also, e-commerce websites enable users to easily submit comments on products and services to improve the quality of products and services, identifying consumer needs, marketing and enhancing inter-company competition [2]. Hence, online reviews are a valid and valuable source of information that plays an important role in shaping the awareness and understanding of customers and users about the characteristics of products and the quality of service [3]. Considering online reviews of products and services by other users or next potential-consumers provide them the opportunity to evaluate and compare products and services before making a purchase decision [3]. It can be claimed that 90% of future sales of products or services depend on the feelings and perceptions contained in the context of previous buyers' opinions. In other words, the purchase likelihood of a product with five reviews texts is 270% greater than the purchase likelihood of a product with no reviews [4]. Therefore, for e-commerce and online sales websites, it is vital that they carefully manage their online reputation using the comments and respond quickly to the complaints and preferences of their users and customers. Accordingly, comment classification and sentiment analysis of the e-commerce websites affect crucially on the survival of the aforementioned e-business in competition with others of its categories. Further, the results can impact the business model, changing and updating warehouses based on customer needs and how services are delivered by e-commerce websites. It will also expand e-commerce and increase customer confidence in online shopping. Manually reviewing the bulk of published comments is difficult, time-consuming, and expensive. On the other hand, people post comments on web pages regardless of correct pronunciation and grammar or they may use emoji or signs. The same problem makes it difficult to extract logical patterns and accurate information from published and so-called unstructured texts [5]. To address these challenges, the sentiment analysis automatically reviews comments published on the web and reports on them in positive, negative and neutral statements [6]. In general, sentiment analysis techniques are implemented at three levels of document, sentence, and feature level, as well as under different approaches, the most important of which are the machine learning-based and lexicon-based approach [7-8].

In recent years, due to the growing importance of a competitive business environment, researchers have sought to use sentiment analysis techniques to classify emotions and summarize opinions about products and services. Because the results not only help other people make better decisions, they are also essential for sellers and producers of products and services, marketing and advertising groups, or other intermediary organizations. In a way, it helps manufacturers and merchants to increase their marketing strategies and understand their customers' needs and interests, as well as to produce and deliver products tailored to their needs and expectations. Accordingly, Cambria believes that "we will gradually come to an era where people's opinions will shape the shape of the final products and services" [9].

This research aims to analyze the sentiment and opinion of Handicraft's consumers since this category of products is of great importance in Iranian e-commerce at both the national and international levels. The analyses were made on evaluating the sentiment comments on the exact "product" and "E-commerce Website Services", separately. To date, there is no evidence on this aspect of this research published in the Persian language. Additionally, the emotion text comments and tweets were classified using XGBoost Algorithm. Besides, the word cloud technique used to display the most important and most frequent words in the studied context. There is no evidence of the application of the last two mentioned techniques in Persian language researches.

2. Research background

In this section, the performed researches about sentiment analysis of products using the machine learning approach are discussed. Although, there is a very limited number of Persian language research in the field of sentiment analysis of products and services, recent and more reliable studies are mentioned here due to the importance of the results provided.

Raeesian Fard and Rezaei [10], presented an aspect-based sentiment analysis model to summarize the text of Persian-language customers' opinions. The data set in their research included customer opinions about the "mobile" product. The proposed model consisted of three parts: 1) preparing the list of features and theoretical words, 2) pre-processing of the text, and 3) determining polarity "semantic orientation" of features. They eventually, displayed the results of summarizing customer reviews in a five-level format (very poor, poor, average, good, very good) to other users.

Haddadi [11], used the supervised algorithms (KNN ، Decision Tree and Naive Bayes) and sentiment analysis techniques to analyze "comparative and non-comparative comments of Persian-language customers about product criticism". The data set in his research included customer reviews that have criticized the "two models of mobile phones at the digikala site". His research results demonstrated that removing "comparative entities" in product reviews can improve the accuracy of the algorithm for classifying comparative comments.

Fan et al. [12] believe that the text of online reviews has a significant impact on product sales. Accordingly, in their research, they have used a machine learning approach and Bass and Norton models to predict product sales based on the context of online reviews. In the beginning, the preprocessing steps were performed on the texts of the comments about the "automotive industry", using word segmentation and word frequency statistics methods. The Naïve Bayes algorithm was then applied for sentiment classification and sentiment analysis. Next, Bass and Norton models were used to predict sales, based on Naïve Bayes' results, which revealed that the Bass model had a better performance in predicting sales. Khan and Malik [13] worked on sentiment classification of Urdu-speaking customers. Initially, pre-processing was performed on 2000 automobiles reviews texts. Next, for the feature selection, they used the “String To Word Vector” method in WEKA software. Finally, for sentiment classification to two positive and negative groups, SVM, KNN, Decision Tree, Multinomial Naïve Bayes, Deep Neural Network, Random Forest, Bagging, AdaBoost algorithms were applied. The results show that Multinomial Naïve Bayes with 89.75% accuracy has the best performance.

Das and Chakraborty [14] used the TF-IDF method with the Next Word Negation (NWN) for sentiment classification. For the feature selection step, they compared the performance of three techniques of the binary bag of words model, the TF-IDF model and the TF-IDF hybrid model with the nearest negative word (NWN). The results show that the combined TF-IDF and NWN models increase the accuracy of the sentiment classification model compared to the other two methods. It was also found that the linear support vector machine algorithm performs better than the Multinomial Naïve Bayes and the combined algorithm of Max Entropy Random, Forest (MERF) has the best performance.

3. Research Implementation Method

Overall, the sentiment analysis framework includes three stages of data preparation (collecting opinions, pre-processing, feature selection), Review Analysis (identifying, extracting useful information from comment texts using different algorithms and techniques), and Sentiment Classification (SC) (Sentiment classification, Evaluation of results) [15-16]. In this study, a framework based on the basic steps of sentiment analysis is suggested, as illustrated in Figure 1, to analyze the users' comments on the e-commerce website.

3-1. Dataset

In sentiment analysis, the collection and selection of appropriate datasets should be based on the type of analysis and concerning for the target subject. Indeed, the richness of the data selected for classification should be considered. For example, to check the sales volume and requirements of the product, it is necessary to collect and review users' opinions regarding the same product and company. So far, most of the existing datasets in sentiment analysis are in English and very few amounts of existing datasets are dedicated to the other languages, i.e. Persian language. However, in this study, web crawling techniques were used through which Persian-language users' opinions were analyzed after being extracted from the e-commerce website and Digikala online sale site.

3-2. Pre-processing

Pre-processing is the first step in matching the textual documents to their presentation in a suitable format, which includes cleaning and preparing the comment text for other processing steps. Pre-processing of Persian language texts is very important due to its specific challenges and complexities [17]. As Persian is one of the non-structured languages, there are some obstacles that researchers encounter when analyzing the context of Persian reviews than other languages [17]. In Persian, if the letters and signs are not written in the same way, the texts cannot be analyzed and the obtained results would not be correct.

Therefore, applying a pre-processing step for Persian language dataset analyses is a pre-cursor of Persian language sentiment analysis.

3-2-1. Normalization

Normalization of textual data has done to sort the text and to integrate the characters by replacing the standard characters in the input text. In Persian, some words have several written forms and some have different structures due to different types

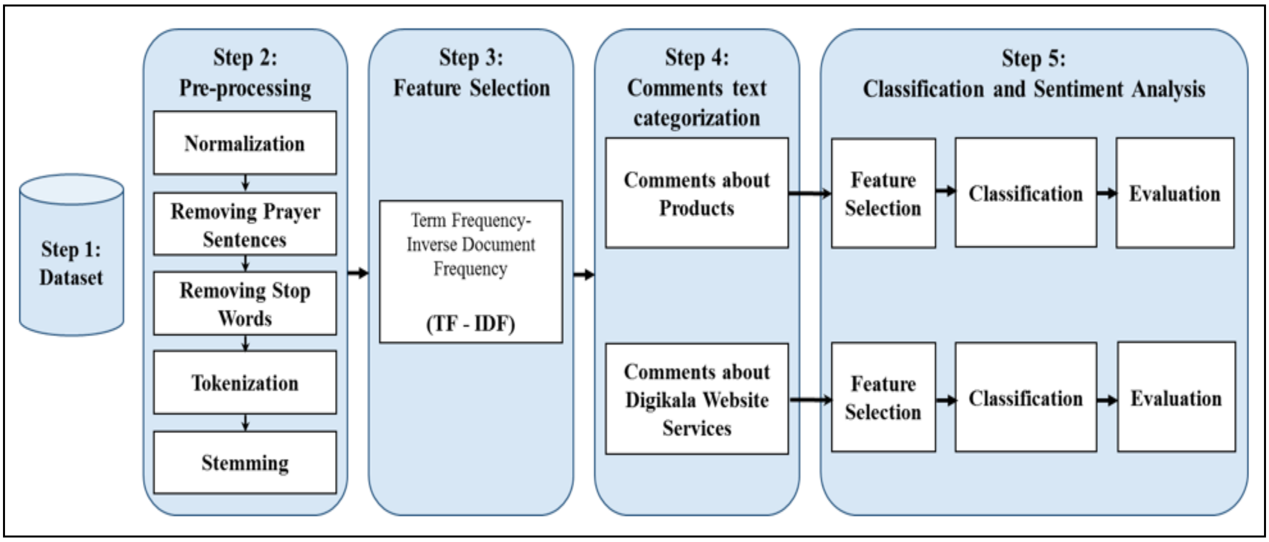


Fig. 1. Proposed Research Architecture

of spaces or half-spaces. Hence, the normalization process in Persian text, punctuation, letters, word spacing, abbreviations, etc. was performed by converting it to a standard form without any semantic changes in the text.

Normalization in Persian text involves converting Arabic characters, also existing figures from different languages to Persian equivalents, aligning distances and converting all forms of distance, half-distance, and no-spaces into a single form, removing symbol "~" in "آ", removing Arabic signs such as Tinnitus, exacerbations and other symptoms [18]. Also, deleting the "-" character, such as converting the word "خــوب" to "خوب" and converting the "ب" attached to the beginning of the words to "به", are other steps in the normalization of the Persian text.

3-2-2. Removing prayer sentences

In the context of the comments, some sentences began with phrases like "ای کاش" (I wish), "امیدوارم" (I hope), "کاش" (wish) they are called "prayer sentences". Using these statements, the speaker does not express his views on the subject but expresses his desire and expected state [10]. It is, therefore, necessary to examine the prayer sentences in the comments and remove them if they lack emotion and value. To the best of my knowledge, there are a few kinds of research in the Persian language which have performed "the removing prayer sentence" step.

3-2-3. Removing stop words

Stop words are part of a sentence that, unlike the many repetitions in the text, are semantically unimportant, of less importance, unnecessary and unused at any stage of the sentiment analysis [19]. Thus removing stop words as trivial words, which are not measured as "keywords" in text, can improve the accuracy and speed of text mining algorithms without destroying the meaning [20]. The list of stop words should be based on the intended use of text mining. For example, the words "هست" (is) and "نیست" (is not) are not important for the subject classification of the text, while they can be effective in emotion analysis. It should be noted, however, that the terms differ depending on the different formats, rules, and commands for language in each country. However, a complete list of stop words in Persian is not provided, which is one of the major challenges in Persian text analyzing research.

3-2-4. Tokenization

Tokenization is the process of breaking a text into meaningful units such as words, paragraphs, symbols, and other elements known as sequences of signs. At this point, punctuation and unnecessary signs will be deleted and word boundary, are determined based on the distance between words. Tokenization often occurs at the word level, and extracted units are used as inputs for other modules such as stemming or tagging. In this study, the text of the comments is broken down into units with the meaning of "word" using the tokenization process. Such as "طراحی روی کیف خاص و سنتی و اندازه مناسب است" ("design on the bag is special and traditional and the size is appropriate.") that is converted to ["طراحی"،"روی"،"کیف"،"خاص"،"و"،"سنتی"،"و"،"اندازه"،"مناسب"،"است"] (["design", "on", "the", "bag", "is", "special", "and", "traditional", "and", "the", "size", "is", "appropriate"]) using tokenization.

3-2-5. Stemming

In each language, words have different appearance forms, depending on their semantic and syntactic roles, indicating their different meanings. In many applications of natural language processing and information retrieval, it is necessary to convert all derivatives of a word to its root (simple word form). Stemming involves removing the suffixes and prefixes of words and extracting their roots according to the rules of word construction [21]. The stemming process allows different forms of a word to be transformed into a single form [22], reducing the number of features and their dimensions. In this research, the deletion of suffixes, prefixes, and root extraction of words under the rooting process are implemented completely on the comments text.

Such as "متاسفم رنگ و اندازه‌های گلدان‌های سفالی خیلی متفاوتند" ("I'm sorry, the colors and sizes of pottery pots are very different.") that is converted to " تاسف، رنگ و اندازه گلدان سفال خیلی متفاوت" ("sorry, the color and size of pottery pot is very different.") using stemming.

3-3. Feature Selection

Semantic comments contain both valuable words and words that have no semantic value. Therefore, to better understand comments and before developing a model for sentiment classification, it is possible to distinguish important words from other words and to identify and select some of the most important features in the context of comments. Therefore, feature selection is considered as one of the most important steps in opinion mining that by eliminating irrelevant and duplicate features, it reduces dimensionality, improves the performance of classifiers, and speeds up machine learning algorithms [23-24].

In this study, the TF-IDF method is used to select appropriate properties of comment text and also to convert text comments to numerical vectors for use in other processing and machine-understandable steps. Hence the Weighted Frequency or Word Frequency-Inverse of Document Frequency (TF-IDF) approach is one of the most common feature weighting methods, which calculates not only the number of repetitions in a document but also the relevance of the word in the document. It assigns weight to each word (feature) in the text and determines how important the word is. The higher the TF-IDF for a word or feature, the more important it is in the text. At the end of the feature selection process using TF-IDF, and based on "The idea TF-IDF is that the most used words, have highest values than the others less used [21]", the words that have the "highest values" or "highest weight" are introduced as "keywords" and "key features" to the next stage for classification.

3-4. Comments Text Classification

In this study, in addition to the sentiment analysis of opinions about products, the emotions in the context of comments related to e-commerce web services are also analyzed. By primarily examining the dataset, it is inferred that some shoppers have examined the characteristics of products such as color, quality, price, and dimensions in their comments. Such as "خوشگله جنسش و کیفیت ساختشم قابل قبوله فقط قیمتش گرونه" ("it is beautiful and its fabric and quality is acceptable but it's expensive ").

Others, in addition to commenting on the product, have also reviewed the services offered by the Digikala website such as delivery time, product packaging and photo and product matching. Like, "کیف خوبیه ولی دیجی‌کالا با تاخیر ارسال کرد" ("It's a good bag, but DigiKala delayed"). A small number of buyers have not reviewed the product features and services offered by the website and have generally commented. Like "عالیه" ("Great" or "awesome"), "افتضاحه" ("Awful") and other comments as expressed. Therefore, at this stage, to analyze the text of comments more accurately and more thoroughly, and to achieve better results in the sentiment classification, the SVM algorithm is used to classify comments into two groups Comments on Products and Comments on Digikala website services.

3-5. Classification and Sentiment Analysis

Sentiment classification focuses on the classification of any text or document and generally involves the process of identifying the emotions or polarity contained in the text or document [25], which classifies texts or documents according to the type of their polarity to their categories of positive, negative or neutral [26]. In this study, the XGBoost algorithm is used to sentiment classification into three groups of positive, negative and neutral, based on machine learning and sentence-level approach. Because studying and classifying neutral opinions plays an important role in the decision making of individuals and organizations, the polarity of neutral emotions is determined only at the sentence level.

3-6. Word Cloud Technique

Word clouds are techniques for visualizing words based on the number of repetitions or the importance of each word in the text [27]. The larger the word in the word cloud is, the greater the weight and frequency of repetition in the text would be, emphasizing its importance. Word cloud results help users quickly identify the most used words in the context of product and service comments and understand their ideas more tangible. Besides, manufacturers, companies and e-commerce websites can use word clouds results to deliver more attractive and effective advertising.

4. IMPLEMENTATION AND ANALYSIS OF RESULTS

In the present study, normalization, tokenization, and stemming were being done using "Hazm" library and removing stop words and removing prayer sentences are done using Programming in Python. Also, for the feature selection process, categorizing the comments' text, "scikit-learn" library was used and finally for the classification and sentiment analysis the "xgboost" library is used in the Python environment.

4-1. Dataset

Initially, 2601 comment's text on handicraft products were selected from the Digikala website as a dataset. In the initial review of the comments which were done manually, the commented text was only "سلام" ("hi"), "با سلام" ("hello") and the duplicate comment text recorded by a user with duplicate content about a product on various dates were deleted. So the dataset contains 2182 comment texts. It was also noted in this review that the data set contains only product buyers' comments. As it can be claimed, all reviews are valid and devoid of any comments provided by some users or even some manufacturers in a false sense with the intent to destroy or promote the product. According to the available dataset, comment providers include 1084 male users, 954 female users, and 144 guest users, which accounted for 49.67%, 43.72% and 6.61% of the total number of comments, respectively. Users have recorded their views on the various categories of the handicrafts found in the Digikala website (Figure 2). Most comments are dedicated to the "کیف چرم" ("Leather Bag") category (577 comments, including 384 male user reviews, 146 female user reviews, and 47 guest user reviews) and the lowest number of comments to the "فیروزه کوبی" ("*Firoozeh koobi*") category (11 comments, including 9 male and 2 female).

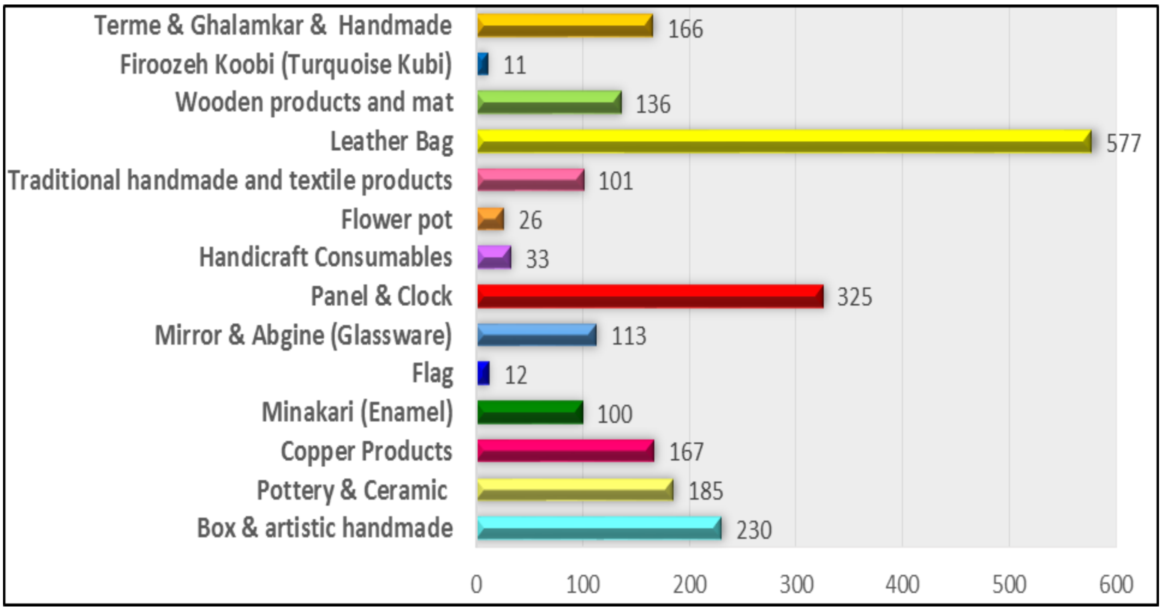


Fig. 2. Number of comments for each products category in the dataset

4-2. Pre-processing

In the initial review of the dataset, it was found that users used different writing methods to record their opinions. Some have recorded their views in the form and structure of conversational sentences, among which there are many misspellings. Others, with a limited number, have submitted their comments by following Farsi grammar rules in the official format. Therefore, before the start of the pre-processing, the text of the informal and folk commentaries was replaced by the process of Slang to Formal Converter and correct forms that correspond to the official textual sculptures, dictionaries and other encyclopedias available in Persian. Also, spelling mistakes and inappropriate words in the comment text were corrected based on the root of the words and the most similar word for the meaningless words whose word or root does not appear in the list of formal or spoken word and phrase words. The normalization process was then performed on 2182 comments. By deleting the prayer sentences the number of comments reached 2175 comments. There were 20828 words in the comments, with the elimination of stop words and product names reaching 11475 words. There were 11475 words in the tokenization process followed by only 5976 in the stemming process.

4-3. Feature Selection

At the end of the feature selection phase, a table is obtained using the TF-IDF method in which the comments text is displayed as numerical vectors and used as input in the next phase. Table 1 lists some of the outputs of the TF-IDF on the comments in this study. A numeric value of zero in each column indicates that the word in question has not been used in the comment text. Before There were 1977 features in the comments before pre-processing execution, however, is reduced to 775 of the total 2175 theoretical examined texts at the end of this phase It is worth to note that during pre-processing of the words, stop words, product names and positive verbs that did not affect the polarity of the comments were eliminated, as well as during the stemming process of different text forms for a word they have been transformed into one. Therefore, the number of features proposed is statistically accurate compared to the number of features pre-processed.

4-4. Comments Text Classification

The SVM algorithm was used to distinguish the comments as either the comments on the products or the comments on Digikala website services. By examining the dataset, it was determined that there were a total of 2502 sentences in the 2175 comment texts, which labeling process was performed manually. The zero-number tag was provided for comments related to the Digkala website services (such as service, shipping, packaging, image and product description). Instead, the one-number tag was provided regarding the numerical labels related to product specification (such as color, price, design, quality, size, and fabric) and general product-related comments (such as excellent, awful, ...) .Then 80% of the comments (2001 comments) were assigned to the training group and 20% (501 comments text) to the experimental group.

After creating the model using SVM, the comments were classified into two groups with 99.2% accuracy. Finally, the performance of the model was evaluated using the k-fold and confusion matrix (Table 2).

In the end, it was identified that using the SVM algorithm, among the 2502 sentences in 2175 comment texts, 2096 comments (sentences) were on products and 406 comments (sentences) were on Digikala website services (Figure 3).

4-5. Classification and Sentiment Analysis

In this step, the XGBoost algorithm is used to analyze and classify the sentiment in the comments into three groups of positive, negative and neutral. It should be noted (it is necessary to note), each of the stages of feature selection, polarization, Parameter Tuning of XGBoost, and the creation of an XGBoost sentiment classification model, are implemented separately on Each group of comments on products and service reviews. The sentiment classification in this study consists of four steps, which are described below:

TABLE 1. DISPLAY TEXT OF COMMENTS AS NUMERICAL VECTORS USING TF-IDF

A	B	C	D	E	F	G	H	I	J	K	L
Comment text	Awful	Size	Good	Happy	Beautiful	Very	Satisfied	Color	Cute	Construction	Pretty
43	0	0	0.311118	0	0	0.539417	0	0	0	0	0.469915
44	0	0	0	0.586917	0	0	0	0	0.221439	0	0.348944
45	0.41541	0	0	0	0	0	0	0	0	0.360966	0
46	0	0	0	0	0.823559	0	0	0	0	0	0
47	0	0	0	0	0	0.267019	0.413437	0.409239	0	0	0
48	0	0	0	0	0	0.99013	0.153306	0.303498	0	0	0
49	0	0.366395	0.253634	0	0	0	0	0	0	0.451909	0.38309
50	0.60131	0	0.293254	0	0	0	0	0	0	0	0
51	0	0	0	0	0	0.435655	0	0	0.481688	0	0
52	0	0	0.282658	0	0	0	0	0	0.270928	0	0
53	0	0.427759	0	0	0.631002	0	0	0	0	0	0
54	0	0	0	0	0.315444	0.256654	0	0	0	0	0
55	0	0	0	0	0.736841	0.299758	0	0	0	0	0
56	0	0	0.168867	0	0	0	0	0.224363	0	0.300877	0.255058
57	0	0.436817	0.302383	0	0	0	0	0	0.289834	0	0
58	0	0.56258	0.194721	0	0	0	0	0	0	0	0
59	0	0	0.376897	0	0	0	0	0	0	0	0
60	0	0	0.563065	0	0	0	0	0	0	0	0
61	0	0.351066	0	0	0	0.210667	0	0	0	0	0

Step 1: Using TF-IDF, the features of each group of comments (products and services) were identified and selected (Table 3).

Step 2: Python language libraries used to determine text polarity do not support Farsi. Thus, the labeling of the polarity of the comments was done manually, and to each comment text in each group of comments (products and services), a polarity was assigned from a set of numbers (+1, 0, -1) that indicate the emotion orientation with "+1" positive polarity, "0" neutral polarity, and "1-" negative polarity, respectively.

Step 3: The XGBoost model in Python for classification, is called the XGBClassifier and contains parameters that use the Grid Search method separately for each group of comments (products or services) before creating a model to obtain the best accuracy in emotion classification.

Step 4: To create the sentiment classification model, each group of comments received 80% and 20% of the training and experiment groups, respectively (Table 3). Finally, the sentiment classification model was created separately based on the optimal parameters associated with each group of comments set and implemented on the comments text, the results are as follows:

Product comments were classified into three groups: positive, negative and neutral, with 95.23% accuracy and the model performance was evaluated using confusion matrix and k-fold (Table 4).

It was found that out of 2096 product reviews, 1447 had positive polarity, 40 had neutral polarity and 609 had negative polarity (Figure 4).

Comments on services were classified into three groups: positive, negative and neutral, with 95.12% accuracy, and model performance was evaluated using confusion and k-fold matrices (Table 5).

It was found that out of 406 comments about services, 153 had positive polarity, 6 had neutral polarity and 247 had negative polarity (Figure 4).

TABLE 2. PERFORMANCE EVALUATION OF SVM

Confusion Matrix		***Predicted values***		
		Label 0	***Label 1***	***Number of comments(501)***
Actual values	***Label 0***	75	2	77
	Label 1	2	422	424
10-fold	***F1-score***	***Recall***	***Precision***	***Accuracy***
99.5	96.8	97	97	99
Model accuracy: 99.2%				

TABLE 3. DATASET DESCRIPTION

Dataset				
Comments	***Number of Comments***	***Number of Features***	***set Train***	***set Test***
Handicraft Products	2096	695	1676	420
Digikala website services	406	203	324	82

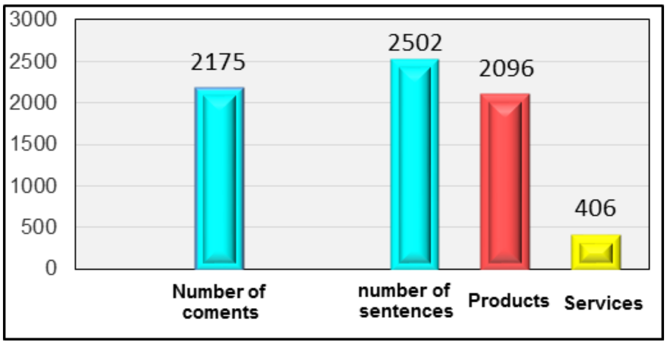


Fig. 3. The total number of comments and sentences recorded on the e-commerce website Digikala. The comments on products and Digikala website services are reported separately.

In the end, the most important words in each group of texts about products and services comments are illustrated using the word cloud technique in Figure 5.

TABLE 4. PERFORMANCE EVALUATION OF XGBOOST (FOR THE COMMENTS ON THE PRODUCTS)

Confusion Matrix		Predicted values			
		Label -1	*Label 0*	*Label 1*	*Number of Comments*
Actual values	*Label -1*	101	0	8	109
	Label 0	1	3	2	6
	Label 1	8	1	296	305
10-fold	*f1-score*	*Recall*	*Precision*	*Accuracy*	*Total Comments*
94.39	83.69	80	88	95	420
Model accuracy: 95.23%					

TABLE 5. PERFORMANCE EVALUATION OF XGBOOST (FOR THE COMMENTS ON DIGIKALA WEBSITE SERVICES)

Confusion Matrix		Predicted values			
		Label -1	*Label 0*	*Label 1*	*Number of Comment*
Actual values	*Label -1*	47	0	2	49
	Label 0	1	0	0	1
	Label 1	1	0	31	32
10-fold	*f1-score*	*Recall*	*Precision*	*Accuracy*	*Total Comments*
95.07	94.7	95	95	95	82
Model accuracy: 95.12%					

5. CONCLUSION

Recent trends in buying products and services from online sale e-commerce websites show that customers and users read reviews of previous buyers before deciding to buy and will be more willing to buy recommended products and services from other buyers than non-recommended products and services. Therefore, opinion mining, by sentiment classification about products/services, seeks to provide important information for other users, companies, and organizations to make appropriate decisions. Also, manufacturers and suppliers of products/services, and e-commerce websites, using opinion mining and its results are trying to identify the patterns of consumer behavior that make a lot of changes in the structure of businesses and using it, take effective steps to succeed.

This research aimed to analyze the content of users' comments on online sale e-commerce websites. The sentiment analysis techniques were used at the sentence level and machine learning approach. A pre-processing step and then the TF-IDF method for selecting features were implemented in the comments text. The comments text' were classified into two groups of products and services comments using the SVM algorithm with 99.2% accuracy. However, the sentiment of comments was classified into three groups of positive, negative and neutral using XGBoost algorithm. The results show 95.23% and 95.12% accuracies for the classification of sentiments in comments about products and about services, respectively. finally, based on the results obtained from the research, it can be acknowledged that sentiment analysis of comment texts about the products and services offered on the e-commerce website is very important and valuable; and in addition to volume of sales of the products/services, it affects e-commerce website reputation which is one of the key parameters in selecting products and services for other users. It is also a cost-effective way to gain other users' trust and free marketing content for the websites and products offered.

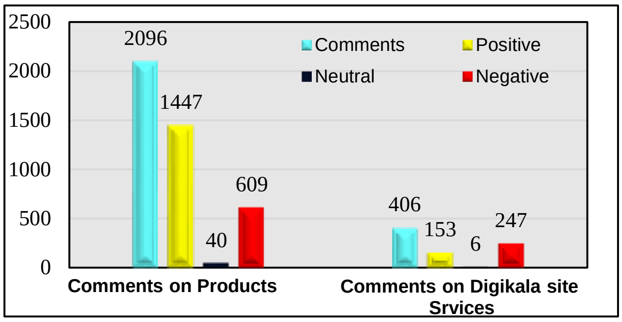
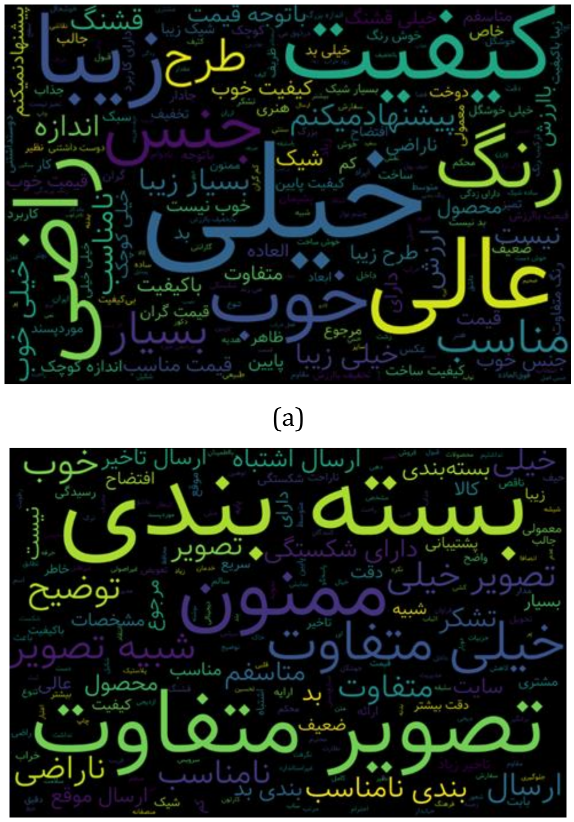


Fig. 4. Sentiment Classification of comments on both products and Digikala site services.



(a)

(b)

Fig. 5. Word cloud. (a) Word cloud for comments on products. (b) Word cloud for comments on services.

REFERENCES

[1] F. Noor, M. Bakhtyar, and U. Baber, "Sentiment Analysis in E-commerce UsingSVM on Roman Urdu Text", In *International Conference for Emerging Technologies in Computing*, Springer, 2019, pp. 213-222.

[2] V. M. Pradhan, J. Vala, and P. Balani, "A Survey on Sentiment Analysis Algorithms for Opinion Mining," *Int. J. Comput. Appl*, vol. 133, no. 9, pp. 7-11, January 2016.

[3] X. Li, C. Wu, and F. Mai, "The effect of online reviews on product sales: A joint sentiment-topic analysis," *Inf. Manag.*, vol. 56, no. 2, pp. 172–184, 2019.

[4] IMC Spiegel Research Center, "How Online Reviews Influence Sales", The Medill IMC Spiegel Research Center, 2017.

[5] R. Irfan, C. K. King, D. Grages, S. Ewen, S. U. Khan, S. A. Madani, ... and N. Tziritas, “A survey on text mining in social networks,” *Knowl. Eng. Rev.*, vol. 30, no. 2, pp. 157–170, 2015.

[6] B. Liu, “Sentiment Analysis and Opinion Mining,” *Lang. Arts Discip.*, no. May, p. 167, 2012.

[7] Y. Fang, H. Tan, and J. Zhang, “Multi-strategy sentiment analysis of consumer reviews based on semantic fuzziness,” *IEEE Access*, vol. 6, pp. 20625–20631, 2018.

[8] M. Khader, A. Awajan, and G. Al-Naymat, “The Effects of Natural Language Processing on Big Data Analysis: Sentiment Analysis Case Study,” *ACIT 2018 - 19th Int. Arab Conf. Inf. Technol.*, Lebanon, Werdanye, IEEE, 2019, pp. 506–513.

[9] E. Cambria and A. Hussain, “Sentic Computing,” Cognit. Comput., vol. 7, no. 2, pp. 183–185, 2015.

[10] M. Raeesian fred and M. Rezaei, “Summarization of Persian language Customers’ opinions about Product Features,” Int. Conf. Electr. Comput. Mech. Iran, Tehran, https//www.civilica.com/Paper-ECMCONF01-ECMCONF01_069., pp. 1–21, 2018. (in persian)

[11] N. Soroush Haddadi, “Analysis, Comparative Opinions on Persian Product Reviews for Competitive Intelligence.,” Congr. New Res. Ideas Eng. Sci. Electr. Comput. Sci. ,Sari, https//www.civilica.com/Paper-SETCO01 SETCO01_132., 2018. (in persian)

[12] Z. P. Fan, Y. J. Che, and Z. Y. Chen, “Product sales forecasting using online reviews and historical sales data: A method combining the Bass model and sentiment analysis,” *J. Bus. Res.*, vol. 74, pp. 90–100, 2017.

[13] M. Khan and K. Malik, “Sentiment classification of customer’s reviews about automobiles in Roman Urdu,” *Adv. Intell. Syst. Comput.*, vol. 887, pp. 630–640, 2018.

[14] B. Das and S. Chakraborty, “An Improved Text Sentiment Classification Model Using TF-IDF and Next Word Negation,” arXiv Prepr. arXiv1806.06407, 2018.

[15] A. Alaei, S. Becken, and B. Stantic, “Sentiment analysis in tourism: Capitalising on Big Data,” *J. Travel Res.* 58, no. 2 175-191., 2019.

[16] A. Sharma and A. Saha, “A Comparative Study of Different Approaches Used For Sentiment,” Available http//dx.doi.org/10.2139/. SSRN 3289248, pp. 1–4, 2018.

[17] M. Basiri, A. nilchi, and N. Ghassem-Aghaee, “A Framework for Sentiment Analysis in Persian,” *Open Trans. Inf. Process.*, vol. 1, no. 3, pp. 1–14, 2014.

[18] H. Azizi and M. Khademi, “Increase Business Intelligence Based on opinions mining in the Persian Reviews Introduction,” *Int. Acad. J. Sci. Eng.*, vol. 2, no. 12, pp. 20–30, 2015.

[19] T. U. Haque, N. N. Saber, and F. M. Shah, “Sentiment analysis on large scale Amazon product reviews,” *2018 IEEE Int. Conf. Innov. Res. Dev. ICIRD*, May 2018, pp. 1–6.

[20] S. Vijayarani, J. Ilamathi, and Nithya, “Preprocessing Techniques for Text Mining - An Overview,” Int. J. Comput. Sci. Commun. Networks, vol. 5, no. 1, pp. 7–16, 2015.

[21] O. R. Llombart, “Using Machine Learning Techniques for Sentiment Analysis `,” Final Proj. Comput. Eng. Sch. Eng(EE). Univ. auto` noma de barcelona (UAB), 2017.

[22] R. Ahuja, A. Chug, S. Kohli, S. Gupta, and P. Ahuja, “The impact of features extraction on the sentiment analysis,” *Procedia Comput. Sci.*, vol. 152, pp. 341–348, 2019.

[23] S. Rohaidah, A. Abu, and M. Ridzwan, “A review of feature selection techniques in sentiment analysis,” vol. 23, pp. 159–189, 2019.

[24] M. Tubishat, M. A. M. Abushariah, N. Idris, and I. Aljarah, “Improved whale optimization algorithm for feature selection in Arabic sentiment analysis,” Appl. Intell., vol. 49, no. 5, pp. 1688–1707, 2019.

[25] D. Lee, O.-R. Jeong, and S. Lee, “Opinion Mining of Customer Feedback Data on the Web,” Proc. *2nd Int. Conf. Ubiquitous Inf. Manag. Commun.*, ACM, 2008, pp. 230-235.

[26] S. Alhojely, “Sentiment Analysis and Opinion Mining: A Survey,” Int. J. Comput. Appl., vol. 150– No.6, pp. 22–25, 2016.

[27] R. Ragini, P. M. R. Anand, and V. Bhaskar, “Big data analytics for disaster response and recovery through sentiment analysis*,” Int. J. Inf. Manage.*, vol. 42, no. September 2017, pp. 13–24, 2018.

Faezeh Forootan was born in 1991 in Tehran, Iran. She received her B.Sc. in Computer Engineering - Software from Islamic Azad University, North Tehran Branch in 2015 and M.Sc. in Information Technology Engineering (E-Commerce) from the University of Ghiaseddin Jamshid Kashani Abyek (Qazvin Province, Iran), in 2019. Her master's thesis is in the field of Sentiment Analysis User Comments On E-commerce Online Sale Websites. Her interests include data mining, text mining, sentiment analysis in text, image and sound, machine learning.

Mohammad Rabiei was born in 1983 in karaj, Iran. He received his B.Sc. in Computer Engineering - Hardware from Islamic Azad University in 2006 and received the M.Sc. degree at the Industrial University of Science and Technology (IUST), Iran, in March 2009 with the highest mark, by discussing a thesis concerning the “Human information literacy and e-readiness”. Also, he received his Ph.D. in Information Technology in Industrial Engineering (Robotics) from Udine University, Italy in 2015 and PhD specializing in ontology in robotics from the University of Leuven, Belgium in 2016. He has been an assistant professor and faculty member at department of Computer Engineering, University of Eyvanekey and also assistant professor at University of Ghiaseddin Jamshid Kashani Abyek since the year 2016 to the present. His research interests include Image processing, machine vision, natural language processing, machine learning, deep learning, Implementing robotics projects and industrial robots using interdisciplinary science and Implementing new business intelligence techniques in public and private organizations, Managing customer relationship and loyalty and e-commerce, Implement social networks in e-marketing and replace this advertising strategy.