

# USING NLP APPROACH FOR ANALYZING CUSTOMER REVIEWS

Saleem Abuleil and Khalid Alsamara

MMIS Department, Chicago State University, Chicago USA

sabuleil@csu.edu

kalsamar@csu.edu

## **ABSTRACT**

*The Web considers one of the main sources of customer opinions and reviews which they are represented in two formats; structured data (numeric ratings) and unstructured data (textual comments). Millions of textual comments about goods and services are posted on the web by customers and every day thousands are added, make it a big challenge to read and understand them to make them a useful structured data for customers and decision makers. Sentiment analysis or Opinion mining is a popular technique for summarizing and analyzing those opinions and reviews. In this paper, we use natural language processing techniques to generate some rules to help us understand customer opinions and reviews (textual comments) written in the Arabic language for the purpose of understanding each one of them and then convert them to a structured data. We use adjectives as a key point to highlight important information in the text then we work around them to tag attributes that describe the subject of the reviews, and we associate them with their values (adjectives).*

## **KEYWORDS**

*Sentiment Analysis, NLP Arabic Language*

## **1. INTRODUCTION**

The web has made it possible for companies to discover what people are saying about their brands online, either in mainstream media like online newspapers and magazines, or on social media. Consumers now search for opinions online before, during, and after a purchase. The next step for brands is finding out whether people are talking positively or negatively about their brand, and why. Some online ratings provide a number but not the reasoning behind it, and may only present half of the story. The process of analyzing user's opinion or sentiment about particular services or products and their features is called opinion mining or sentiment analysis. There are two major approaches for performing sentiment analysis; statistical model based approaches and Natural Language Processing (NLP) based approaches to creating rules.

With the quick growing of e-commerce, the number of products sold on the web, the number of services offered on the web and number of online buyers and shoppers increased dramatically. It became an important task for online merchants to enable their customers to review or to express opinions on the products that they have purchased and use this feature to enhance customer

satisfaction about their products and services and also use it as a marketing tool. Almost all firms who sell goods and products on the web make it part of their business to collect and gather information about their services and goods they provide to their customers, in addition to many independent companies who collect customer opinions. Analyzing this information, summarize it, and make it available for decision makers to observe how consumers think about products and services, make it available as well for customers to assist them to make a comparison to improve their decisions before they make any order or request any service.

## **2. BACKGROUND AND RELATED WORK**

The idea of opinion mining or sentiment analysis is to process a set of search results for a given entity, generating a list of attributes which are termed as opinion features of that entity. As a result of increasing number of people who are writing reviews on the Web, the number of reviews for products and receives grows rapidly. Some popular products can get hundreds of reviews at some large merchant sites, some reviews are short and easy to read and decision about them, but some reviews are long and have only a few sentences containing opinions on the product, that makes it hard for a potential customer to read them to make decision on them. A large number of reviews also makes it hard for product manufacturers to keep track of customer opinions of their products.

The extraction of a sentiment can be made either on a whole document (document-level SA), on each paragraph (paragraph-level SA), or on each sentence (sentence-level SA) [11]. Zen Hai and C Yang [17] proposed a method to identify opinion features from online reviews by exploiting the difference in opinion feature statistics across two corpora, one domain-specific corpus and one domain-independent corpus, this is captured by a measure called Domain relevance. They first extracted a list of candidate opinion features from the domain review corpus by defining a set of syntactic dependence rules. For each extracted candidate feature, they then estimated its intrinsic-domain relevance (IDR) and extrinsic-domain relevance (EDR) scores on the domain-independent and domain-independent corpora, respectively. These values are compared with a threshold and are identified as best candidate features. Vasileios Hatzivassiloglou and Jance Wiebe [14] study the effects of dynamic adjectives, semantically oriented adjectives, and gradable adjectives on a simple subjectivity classifier, and establish that they are strong predictors of subjectivity. They have proposed a method for predicting subjectivity of opinions at sentence level by a supervised classification method. A trainable method that statistically combines two indicators of gradability is presented and evaluated, complementing existing automatic techniques for assigning orientation labels. Pang and Lee [3] proposed a machine-learning method that applies text-categorization techniques to just the subjective portions of the document to determine sentiment polarity. They examined the relation between subjectivity detection and polarity classification, showing that subjectivity detection can compress reviews into much shorter extracts that still retain polarity information at a level comparable to that of the full review. they have also shown that employing the minimum-cut framework results in the development of efficient algorithms for sentiment analysis.

Ryan McDonald and Kerry Hannan [13] have investigated the use of a global structured model that learns to predict sentiment on different levels of granularity for a text. The proposed model has the advantage of building the single model for all granularity levels. Labeling is done by MIRA algorithm which works at document and sentence level by applying a weight vector to each label. They showed that this model obtains higher accuracy than classifiers trained in

isolation as well as cascaded systems that pass information from one level to another at test time. Lizhen Qu and Georgiana Ifrim [10] have proposed a set of techniques for mining and summarizing product reviews based on data mining and natural language processing methods by performing three steps: mining product features that have been commented on by customers; identifying opinion sentences in each review and deciding whether each opinion sentence is positive or negative; summarizing the results.

Yessenalina and Cardie [1] Have presented a matrix-space model for ordinal scale sentiment prediction and an algorithm for learning such a model. The proposed 180 model learns a matrix for each word; the composition of words is modeled as iterated matrix multiplication. In the context of the phrase-level sentiment analysis task, their experimental results show statistically significant improvements in performance over a bag-of-words mode. Wei Jin and Hung Hay Ho [15] proposed a model that provides solutions for server problems that have been not provided by previous approaches. This system can self-learn new vocabularies based on the pattern it has learned, which is used in text and web mining. A novel approach is used to handle situations in which collecting a large training set could be expensive and difficult to accomplices. Guang Qiu, Bing Liu, Jiajun Bu and Chun Chen [8] have emphasized on two important tasks in opinion mining, namely, opinion lexicon expansion and target extraction. they proposed a propagation approach to extract opinion words and targets iteratively given only a seed opinion lexicon of small size. The extraction is performed using identified relations between opinion words and targets, and also opinion words/targets themselves. Bo Pang and Lillian Lee [6] examine the relation between subjectivity detection and polarity classification. The subjectivity detection can compress reviews in shorter extracts that still retains polarity information at a level comparable to that of the full review. By using Naive Bayes polarity classifier the subjectivity extract are shown to be more effective input than the originating document. They show that the minimum-cut framework results in the development of an efficient algorithm for sentiment analysis. Via this framework, contextual information can lead to statistically significant improvement in polarity classification accuracy. Niklas Jacob and Iryna Gurevych [12] have shown how a CRF-based approach for opinion target extraction performs in a single- and cross-domain setting. They have presented a comparative evaluation of our approach on datasets from four different domains.

### 3. OUR CONTRIBUTION

Adjectives play a key role in this paper, they represent values of attributes and features of products and services. *In linguistics, an adjective is a describing word, the main syntactic role of which is to qualify a noun or noun phrase, giving more information about the object signified*<sup>1</sup>. Adjectives are one of the Arabic parts of speech. Arabic Adjectives are words that describe or modify another person or thing in the sentence. In Arabic adjectives are of the form فعيل Fa3iil, like كبير kabiir big, صغير saghriir small. One rule is that if a noun is definite the adjective has to be definite, like in البيت الكبير Al-bait Al-kabir The House The Big (Al is the Arabic indefinite article). Just like Spanish & German, Arabic has masculine and feminine adjective forms, in Arabic to form a feminine adjective from the masculine, you simply add “taa’ marbuta” which looks like (ة, ة) to the end of the adjective for example (he) Beautiful Jameel جميل (masculine) and (she) Beautiful jameela جميلة (feminine). *In Classical Arabic, adjectives must agree with the nouns they modify in terms of gender (masculine or feminine),*

---

<sup>1</sup> Wikipedia website, the free encyclobidia, “<https://en.wikipedia.org/wiki/Adjective>”

*number (singular, dual or plural), grammatical case (subject, direct object or prepositional) and state of definiteness (whether the noun is definite or indefinite)<sup>2</sup>.*

In this paper we study customer opinions (reviews) written in the Arabic language for the purpose of understanding each one of them and then convert unstructured text to a structured data, very little work has been done in this area in the Arabic language and there is big need to contribute to it. We emphasize in this paper on two main elements: attribute and attribute value. For example TV product described by some attributes such as a screen, sound, price, size, where each attribute has a certain value such as good, bad, high, low, beautiful. Attributes are two types either simple or compound, simple attribute consists of one word such as sound, price, and size, compound attributes consists of two words to emphasise a specific feature such as sound *quality*, picture *quality*, resolution *accuracy*, and sound *clearness*, compound attribute comes in three main formats: *اللون الشاشا* *colors of screen*, *الشاشا الوانه* *screen colors*, *الشاشا ذات الولا* *screen with colors*. In some cases adjective is attached to a special word to neglect it; change the status from positive to negative; such as not i.e. *not good*.

Unlike English, Arabic adjectives follow the noun they modify, which is somehow easier, because when you start with the noun first you will easily modify the adjective that comes afterward accordingly either to its masculine, feminine, dual or plural form. The noun in this context is the attribute described by the adjective. After studying hundreds of reviews we came up with novel approach consists of three steps as it shown in figure 1 to understand customer reviews written in the Arabic language.

To support our approach and to achieve our goal we collect attributes and adjectives and classify new adjectives while we are running our approach and save them in two main tables: attributes table and adjectives table, attributes tables include both simple attributes and compound attributes, each entry in this table has a pair of two roots represent a certain attribute, for simple attributes the second root is null. Adjectives table includes root of each adjectives and its classification either good or bad, we also have collected neglect tools (words) and saved them in a list.

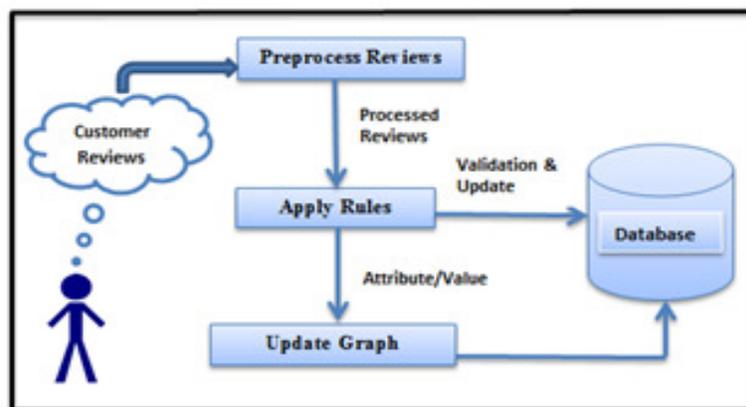


Fig 1: Approach Structure

Our approach consists of three steps as following:

<sup>2</sup> Learn Languages with Speak7 website, "<http://arabic.speak7.com>"

- 1- Preprocess Reviews: read reviews, use a morphology and part-of-speech tagging systems to:
  - a. Find part-of-speech and root for each word in the text
  - b. Identify adjectives in the text
  - c. Check if neglected tool (word) is attached to the adjectives
- 2- Apply Rules: Extract attributes and associate them with their values (adjectives) that are labeled in step #1.
  - a. Tag up to two words headed by an adjective, stop when encountering a verb, particle or punctuation mark.
  - b. Use the following rules to form adjective phrases:
 
$$\text{Adjective Phrase} \rightarrow \langle \text{Attribute} \rangle \langle \text{Adjective} \rangle$$

$$\quad \quad \quad | \langle \text{Attribute} \rangle \langle \text{Neglect-Tool} \rangle \langle \text{Adjective} \rangle$$

$$\text{Attribute} \rightarrow \text{Simple Attribute} | \text{Compound Attribute}$$
  - c. Check if <adjective> is already in adjectives table, find its classification, either positive or negative, otherwise classify it and update the adjectives table
  - d. Check if <attribute> either if it is a simple or compound is in attributes table, if not validate it and update attributes table
- 3- Update Graph: use the output from step #2 (attributes/values) to update graph by updating frequency of each node and each edge. Each node in the graph contains either an attribute or a value, attribute nodes connected to values nodes through edges as shown in figure 2.

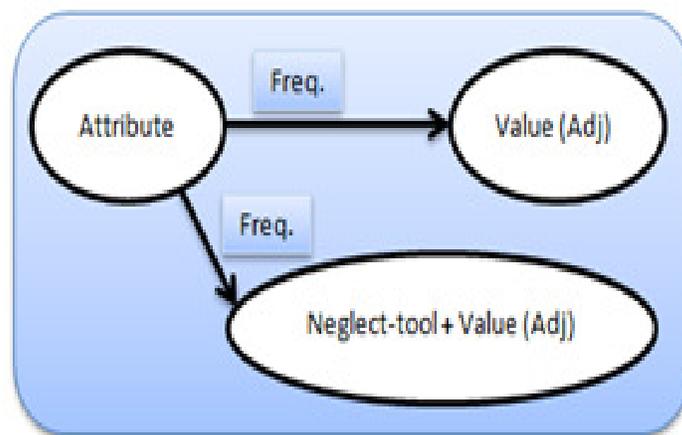


Figure 2. Graph Nodes

#### 4. ANALYSIS

In the following example, we demonstrate how we use our approach to convert customer reviews from unstructured text to a structured data. The reviews we use in this example are about Samsung LED 4009MS-U7D 40 inch TV posted on egypt.souq.com website. First, we run a morphology and part-of-speech tagging systems to identify adjectives and to find part-of-speech and root for each word in the text. The following is a sample review shows just adjectives, root and part-of-speech of each word are not shown.

احببت تلفزيون سامسونج كثيرا, استطيع ان استخدمه لليوتيوب وتطبيقات اخرى حاولت الطباعة باللغة العربية لبعض التطبيقات ولم استطع ذلك استخدمه للاكس بوكس و العب عليه طوال الوقت. صورة نقية وصوت واضح لكن الظهر الشاتنه ليس قوي. لا يوجد كابل ستريو للسماعات الخارجيه وكنت مندهشا عندما وجدت ان الجهاز ياتي معه ايضا محول اضافي

*I like Samsung TV a lot; I can use YouTube and other applications I use for Xbox as well as other games most of the time. Pure picture, clear sound but screen back is not durable, I tried to type in Arabic while I am using different applications but I could not. There is no cable for external stereo headphones. Also, I was surprised when I have seen that the TV comes with external adapter*

Second, we tag up to two words headed by adjective, stop when encounter a verb, particle or punctuation mark, we apply some rules to form adjective phrases, check category of each adjective either positive or negative, identify attributes and associate them with their values, validate and update adjectives and attributes tables, the output of the second step is three adjective phrases as follows:

<Simple Attribute: صورة picture> <Value (positive): نقية pure>

<Simple Attribute: صوت sound> <Value (positive): واضح clear>

<Compound Attribute: ظهر شاتنه screen back> <Neglect-Tool: ليس not> <Value (positive): قوي durable>

Third, we update the graph. Table 1 shows the result of 100 customer reviews.

In the above example, we found 70% of customers gave positive review for TV sound, 30% gave bad review, while 80% of customers gave positive review for TV screen and 20% gave negative review and just 8% gave positive review for the screen back and 92% gave negative review about it. Attribute frequency: TV sound repeated 50 times, TV picture 75 times and TV screen back 60 times, this gives the indication of the importance of each attribute in the reviews. In this paper we handled one side of customer reviews that when an adjectives present to describe attributes, another side needs to be studied is when customer reviews mention certain features or attributes without using adjectives to describe them such as:

*I can use it to access YouTube and other applications  
There is no cable for external stereo headphones*

*استطيع ان استخدمه لليوتيوب وتطبيقات اخرى  
لا يوجد كابل ستريو للسماعات الخارجيه*

In the above examples customers providing important facts about the TV, but because of the absent of any adjectives in the text we cannot catch them. One way is to look for some special phrases and work around them such as “استطيع ان استخدمه لل” “I can use it to” and “لا يوجد” “There is no”.

Table 1: Customer Reviews for Samsung LED 4009MS-U7D 40 inch TV

Attribute	Value	Frequency	Positive	Negative
صوت Sound	رائع nice	5	5	0
صوت Sound	عالي high	5	5	0
صوت Sound	نقي pure	5	5	0
صوت Sound	جيد good	10	10	0
صوت Sound	ممتاز excellent	5	5	0
صوت Sound	واضح clear	5	5	0
صوت Sound	رديئ bad	10	0	10
صوت Sound	غير واضح not clear	5	0	5
<b>Total</b>		<b>50</b>	<b>35 / 70%</b>	<b>15 / 30%</b>
صورة Picture	رائع nice	15	15	0
صورة Picture	نقي pure	5	5	0
صورة Picture	جيد good	5	5	0
صورة Picture	ممتاز excellent	10	10	0
صورة Picture	واضح clear	25	25	0
صورة Picture	نغمسه noise	15	0	15
<b>Total</b>		<b>75</b>	<b>60 / 80%</b>	<b>15 / 20%</b>
ظهر الشاشة Screen Back	ليس قوي not durable	5	0	5
ظهر الشاشة Screen Back	غير مقوي not durable	5	0	5
ظهر الشاشة Screen Back	خفيفة light	5	0	5
ظهر الشاشة Screen Back	تعبان bad	10	0	10
ظهر الشاشة Screen Back	ليس جيد not good	15	0	15
ظهر الشاشة Screen Back	ممتاز excellent	5	5	0
ظهر الشاشة Screen Back	ضعيف weak	15	0	15
<b>Total</b>		<b>60</b>	<b>5 / 8%</b>	<b>55 / 92%</b>

## 5. CONCLUSION

In this paper, we have introduced our approach for using NLP to generate some rules to help us understand customer opinions and reviews (textual comments) written in the Arabic language for the purpose of understanding each one of them and convert them to a structured data. In future research we are going to study more reviews from different resources to test our approach on more data and generate more detailed analysis, we are going also to study cases when the adjective is absent, how to analyze text and understand it by looking for certain keywords in the reviews and work around them.

## REFERENCES

- [1] A. Yessenalina and C. Cardie, "Compositional Matrix-Space Models for Sentiment Analysis", Proc. Conf. Empirical Methods in Natural Language Processing, pp. 172-182, 2011.
- [2] B. Liu, "Sentiment Analysis and Opinion Mining", Synthesis Lectures on Human Language Technologies, vol. 5, no. 1, pp. 1-167, May 2012 .
- [3] B. Pang and L. Lee, "A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts", Proc. 42nd Ann. Meeting on Assoc. for Computational Linguistics, 2004.
- [4] B Liu, "Sentiment Analysis and Opinion Mining", Synthesis Lectures on Human Language Technologies, vol.5,no.1, pp.1-167,May 2012.

- [5] E. Cambria, D. Osher and K.Kwok, “Sentic Activation : A two Level Affective Common Sense Reasoning Framework”, Proc.26th AAAI Conf. Artificial Intelligence, pp.186-192, 2012.
- [6] Forman, B. Pang and L. Lee, “A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts”, Proc. 42nd Ann. Meeting on Assoc. for Computational Linguistics.
- [7] G.Qiu , C.Wang, J.Bu , K.Liu and C.Chen, “Incorporate the Syntactic Knowledge in Opinion Mining in User Generated Content”, Proc. WWW 2008 Workshop NLP Challenges in the information Explosion Era, 2008.
- [8] G. Qiu, B. Liu, J. Bu, and C. Chen, “Opinion Word Expansion and Target Extraction through Double Propagation”, Computational Linguistics, vol. 37, pp. 9-27, 2011..
- [9] L. Qu, G. Ifrim, and G. Weikum, “The Bag-of-Opinions Method for Review Rating Prediction from Sparse Text Patterns”, Proc. 23rd Int’l Conf. Computational Linguistics, pp. 913-921, 2010.
- [10] M. Hu and B.Liu, “Mining and Summarizing Customer Reviews”, Proc. 10th ACM SIGKDD Int’l Conf. Knowledge Discovery and Data Mining, pp. 168-177,2004.
- [11] M. Korayem, D. Crandall, and M. Abdul-Mageed. Subjectivity and sentiment analysis of arabic: A survey. In AboulElla Hassanien, Abdel-BadeehM. Salem, Rabie Ramadan, and Tai-hoon Kim, editors, *Advanced Machine Learning Technologies and Applications*, volume 322 of *Communications in Computer and Information Science*, pages 128–139. Springer Berlin Heidelberg, 2012.
- [12] N. Jakob and I. Gurevych, “Extracting Opinion Targets in a Single and Cross-Domain Setting with Conditional Random Fields”, Proc. Conf. Empirical Methods in Natural Language Processing, pp. 1035-1045, 2010.
- [13] R. Mcdonald, K. Hannan, T. Neylon, M. Wells, and J. Reynar, “Structured Models for Fine-to-Coarse Sentiment Analysis”, Proc. 45th Ann. Meeting of the Assoc. of Computational Linguistics, pp. 432-439, 2007.
- [14] V. Hatzivassiloglou and J.M. Wiebe, “Effects of Adjective Orientation and Gradability on Sentence Subjectivity”, Proc. 18th Conf. Computational Linguistics, pp. 299-305, 2000.
- [15] W. Jin and H.H. Ho, “A Novel Lexicalized HMM-Based Learning Framework for Web Opinion Mining”, Proc. 26th Ann. Int’l Conf. Machine Learning, pp. 465-472, 2009.
- [16] Y. Jo and A.H. Oh, “Aspect and Sentiment Unification Model for Online Review Analysis”, Proc. Fourth ACM Int’l Conf. Web Search and Data Mining, pp. 815-824, 2011.
- [17] Zhen Hai, Kuiyu Chang, Jung-Jae Kim, and Christopher C. Yang “Identifying Features in Opinion Mining via Intrinsic and Extrinsic Domain Relevance”, IEEE transactions on knowledge and data engineering, Vol. 26, NO. 3, MARCH 2014.