Rule-based Emotion AI in Arabic Customer Review

Mohamed M.Abbassy[1]
Information Technology Department
Faculty of Computers and Artificial Intelligence
Beni-Suef University, Beni-Suef, Egypt

Ayman Abo-Alnadr[2]
Information System Department
Higher Institute of Management and Information
Technology, Kafer el Shekh, Egypt

Abstract—The e-commerce emotion analysis is notable and the most pivotal advance since it catches the customer emotion in a product, and emotions with respect to product to decide if the customer attitude is negative, positive, or neutral. Posting on the customer's reviews have turned into an undeniably famous path for individuals to share with different customers their emotion and feelings toward a product. This review has a significant impact on sales in the future. The proposed system utilizes mixed word from an adjective (adj) and adverb to improve the emotion analysis process utilized a rule-based emotion analysis. The system extracts an Arabic customer review and computes the frequency of each word. At that point, it computes the emotion and score of each customer review. The system likewise computes the emotion and score of straightforward Arabic sentence.

Keywords—Component; rule-based; emotion; customer review; Arabic

I. INTRODUCTION

With the rapid development of web applications, social network and online shopping, there moved toward becoming audits, comments and feedback generated by customers. These emotions can be about essentially anything, including products, politics issues, news and service. All of which should be handled and broke down to get a good estimation of what the customer thinks and feels. Before the accessibility of automatic emotion analysis tools, the way toward getting customer surveys was an incredibly cumbersome and time-consuming task [1].

Numerous emotion AI were created for English, yet in this paper are attempting to break another ground in this field and concoct a high accuracy Arabic based emotion analysis tool which isn't influenced by the utilization of vernaculars; a tool that enables Arab customers to analyze the e-commerce shopping, enabling them to know the general feeling about products being talked about. The Arabic language has numerous lingos that ought to be considered, wherein every vernacular implications of words can be very surprising. Arabic is a morphologically rich language and this can raise issues for any programmed content examination instrument [2, 3].

The enormous increment in e-commerce shopping in middle east especial Egypt, gulf countries, that made customer reviews significant in decision making procedure of a customer. The quantity of reviews for a product can be very high, particularly for a most prevalent product. A significant number of customers is interested on emotion of a product, so for this reason, they should initially read all the reviews to reach a resolution. What's more, since perusing countless is a dreary procedure and may create upsets in basic decision making [4].

In this way, an effective method for showing the general emotion of a product dependent on customer reviews is required. This paper inquire about in the investigation of product reviews are worried about ordering the general emotion for a specific product. As a customer review does not have a standard structure and may incorporate spelling blunders and equivalent words for the product features, emotion classification per feature can be troublesome.

Emotion AI is a procedure of extracting information from users' assessments. The decisions of the people get influenced by the conclusions of other individuals. Today, if any person needs to purchase an item then the person will initially look through the surveys and emotion about that item via an online shopping, a social network like Twitter, Facebook, and other user forums, at that point recognizable proof of assumption, turns out to be extremely troublesome from this colossal information physically. Thus, there is a need for a computerized emotion analysis system. The fundamental goal of this paper is to perform emotion AI for Arabic sentences.

II. RELATED WORK

Elhawary and Elfeky [5] utilized that gathering Arabic business reviews, and dedicating 80% of the gathered business reviews to prepare their classifier which is utilized to recognize review's records. They developed various Arabic vocabularies used to investigate distinctive Arabic reviews and emotion. The extremity of every Arabic business review whether it is: positive, negative, neutral or mixed is made a decision about dependent on the assembled dictionaries.

Diverse strategies were utilized by El-Halees [6] to decide the extremity of various Arabic `s. The extremity of the entire Arabic comment is resolved first utilizing the vocabulary-based technique, where the output from the primary strategy (dictionary based) is considered as a preparation set for greatest entropy strategy, which is utilized to order these comments.

Another methodology has been proposed depends on translating the source Arabic emotions into English and after that utilization the equivalent relevant procedures to examine the came about English emotions. Almas and Ahmad in [7] utilized machine translation systems to translate the source comment or review from Arabic to English language before passing them to an English based emotion analysis system. The issue of this methodology was the loss of nuance after translating the source to English.

Rushdi-Saleh et al. [8] utilized another methodology was machine learning algorithm to arrange the extremity of Arabic reviews extricated from specific Web pages identified with motion movies. Inui et al. [9] think about receive making an interpretation of suppositions from English to Japanese, trailed by emotion analysis. They applied sentiment-oriented sentence filtering strategy to alleviate numerous interpretation mistakes that happen as a reaction of interpretation to decrease the impact of interpretation blunders in multilingual comment level review.

Choi et al. [10] presents a structure for emotion analysis, focus around the feeling piece of information that is identified with a supposition theme, for example, company or individual. They utilize a domain-specific sentiment classifier for every domain with the recently totaled signs (for example a subject or the theme of the emotion) in light of a proposed semi-supervised strategy. Yi et al. [11], Kim et al. [12], Choi et al. [10] extricate emotion about a subject spotlight on the estimation piece of information that is identified with a conclusion theme. This is characterized as an essential subject of supposition articulation in a sentence, for example, organization, individual or occasion.

III. MATERIAL AND METHODS

To perform emotion AI, basic Arabic content record, tweets or comment in online shopping are inputted by the client. At that point, the system takes a shot at it and figures its emotion and score. The design appeared in Fig. 1 show the working of rule-based emotion AI system.

A. Tokenizer

The system takes a product review in the Arabic language as an input; the input sentence is part into tokens through tokenizer. A token is a piece of an arrangement of characters in content that are combined together as an important semantic unit for handling. The tokenizer changes over a sentence into word level tokens comprising of a word, accentuation marks, and different symbols.
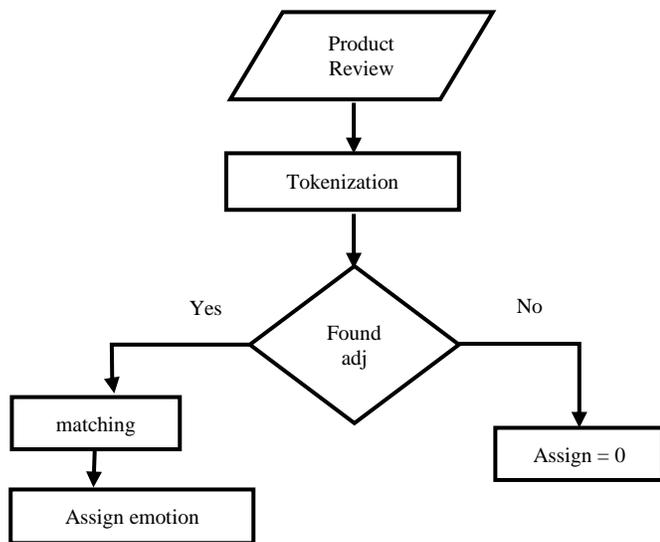

Fig. 1. Proposed Architecture of Rule-based Emotion AI System.

B. Procedure of Final Emotion and Score

Most of the work use adjectives only for emotion analysis, and some of them use nouns, verbs, adverbs or a combination of them. This proposed model use emotion adj and nouns of adj words because that some emotion do exclude any adj, however, express a negative or positive sentiment such as; The vast majority of the work use adjectives just for emotion investigation, and some of them use noun, verb, adverbs or a mix of them. utilize emotion adjectives and noun of adjectives since they noticed that some emotion statements do exclude any adjectives, however, express negative or positive emotion, for example, "حلو مرة" ," جميل جدا" ," ممتاز" ,"جامد جدا" ," رائع". That means "Wonderful", "Excellent", "Very beautiful", "Very nice". All of them are adjectives in this example, although it expresses positive emotions.

IV. APPLYING RULE

An algorithm is proposed to extend and recognize the emotion AI automatic of new feeling words utilizing total procedure and free online Arabic word references and vocabularies and calculate the frequency of each emotion word from a dataset. The scores of adj are appointed between -1 and 1. In the event that any of the token matches with terms of adj and adverb score content then their comparing scores are processed. At that point, the last emotion is assigned to content as positive, negative or natural utilizing calculation has been proposed below. Be that as it may, if any of the tokens in the content does not match with terms of adj score. At that point, 0 scores are assigned to that content lastly no emotion is assigned to that content. For example: "هذا المنتج جميل جدا". In this example, the process of this content will be split into tokens as "هذا", "المنتج" ," جميل", "جدا". Presently, these tokens are coordinated with adj and adverb score content. Here, "جميل", "جدا" are found in adverb and adj word score content separately. Along these, relating scores of "جميل", "جدا" are appointed as 0.1 and 1 individually. And afterward, the last score and emotion are registered by a system utilizing the proposed model in this paper.

A. Role of Score Adjective and Adverb

The score document contains the rundown of pre-processed emotion scores as appeared in the table below. Each line of the scoring document contains a word or expression alongside its emotion score. In the event that a word or expression which is found in content yet not found in score document at that point word or expression is given an emotion score of 0.

At that point, the emotion of content based on assumption scores of the terms in the content is registered. The emotion of content is equivalent to the whole of the assumption scores for each term in the content.

Grammatical forms information is most generally used in linguistic tasks. It is used to disambiguate sense which subsequently is used to coordinate component decision [13]. Researchers basically use adjective (adj) words and adverb as highlights to discover the emotion in content. Adjective (adj) words are most normally utilized as highlights among all grammatical features. There is a solid connection amongst adj and subjectivity of content. Indeed, even every one of the grammatical features assumes a critical role, yet just adj words

as highlights feature the emotions with high exactness. An exactness of around 82.8% has been accomplished in film survey spaces by utilizing adj words just as highlights [14]. A few instances of positive and negatives have appeared in Table I.

B. Role of Score Content

1) Calculate strong function: If the score of an adj is greater than 0 then adj is positive and adv has a place with positive. For example, "جميل" has meant in English "beautiful" is a positive adj and "جداً" mean in English "Very" has a place with a strong adverb.

2) Calculate weak function: If the score of an adj is less than 0 then is a negative emotion and adverb has a place with is negative then score of both adj and adverb will be less than 0. For example, "سيئ" that means in English "bad" is negative adj and "جدا" that mean in English "very" is also a negative adverb. The model has been proposed to calculate strong and weak functions in this work as shown in Fig. 2.

TABLE. I.   SOME WORD OF POSITIVE AND NEGATIVE ADJECTIVE

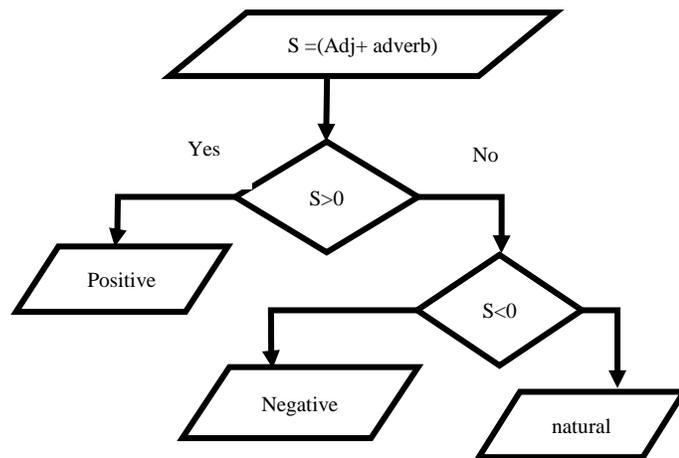| Positive adjective					Negative adjective			
جميل	رائع	جيد	عالى	ماركة	ردئ	سئ	محلى	مستهلك
ممتاز	مبهر	ناجح	اقتصادى	فل	صعب	بطئ	اوفر	عادى
اجمل	مميز	مذهل	صاروخ	فل الفل	ممل	مقزز	تقليدى	زفت
تحفة	حلو	خامة	ملفت	ياعينى	وحش	غالي	صينى	ز الزفت
راض	كويس	قوى	عالمى	ياجمالوا	سلبي	بينة	تقليد	اطران
ممتع	انبسطت	عملى	براند	ياحلوته	فاشل	كوبى	ضعيف	هباب

Fig. 2.   The Rule-Based Model is to Calculate the Score of Arabic Sentences Whether Strong or Weak.

V. DESCRIPTION OF THE PROPOSED ALGORITHM

First step is to take an Arabic sentence as output for a system, the second step is the identification of relations and widespread words, third step is assigning a score to emotion according to pre-processed emotion scores as appeared in table previous, fourth step calculation a final score according to degree of strong or weak function according to relation identification as shown below.

If score > 0

Then Arabic sentence show positive emotion,

else if score < 0:

then the Arabic sentence show negative emotion,

else: the Arabic sentence shows no emotion or normal emotion.

VI. CONCLUSION

In this paper proposed technique to extract emotion focus from Arabic customer review from online shopping. In this paper, proposed Adjective and Adverb together are considered for performing emotion analysis as it gives preferable result. The purposed system helps in computing the emotion and score of customer review.

The proposed system computes the emotion and score of basic Arabic sentences. These sentences are part into tokens by tokenization process. At that point, with the assistance of adjective and adverb score records; emotion and score of these sentences are found by creating database for both negative and positive word used by customer.

REFERENCES

[1] Waters, John K. The Everything Guide to Social Media: All you need to know about participating in today's most popular online communities. Adams Media, 2010.

[2] Chiang, David et al. "Parsing Arabic dialects." Proceedings of the European Chapter of ACL (EACL) 2006: 112

[3] "Morphology (linguistics)", Wikipedia, the free encyclopedia." Last Accessed, 4 May 2019 .http://en.wikipedia.org/wiki/Morphology _(linguistics)

[4] S.Pednekar,K.Patil,R. Sawant,T.Shah, "Sentiment Analysis On Online Product Reviews", International Education & Research Journal [IERJ], pp.130-131,Mar 2017.

[5] Elhaway, Mohamed, and Mohamed Elfeky. "Mining Arabic Business Reviews." Data Mining Workshops (ICDMW), 2010 IEEE International Conference on 13 Dec. 2010: 1108-1113.

[6] A. El-Halees, "Arabic Opinion Mining Using Combined Classification Approach," In: Proceedings of the International Arab Conference on Information Technology, Zarqa, Jordan, 2011.

[7]   Y. Almas,K. Ahmad, "A note on extracting 'sentiments' in financial news in English, Arabic & Urdu," In: The Second Workshop on Computation, al Approaches to Arabic Script-based Languages, Linguistic Society of America 2007 Linguistic Institute, Stanford University, Stanford, California., Linguistic Society of America, pp. 1–12, 2007.

[8]   M. Rushdi-Saleh, M. Teresa Martín-Valdivia, L. Alfonso UreñaLópez,J. M. Perea-Ortega, "Bilingual Experiments with an ArabicEnglish Corpus for Opinion Mining. Language," In: Proceedings of Recent Advances in Natural Language Processing, Hissar, Bulgaria. pp. 740-745, 2011.

[9]   T. Inui, M. Yamamoto "Applying Sentiment-oriented Sentence Filtering to Multilingual Review Classification," In: Proceedings of the Workshop on Sentiment Analysis where AI meets Psychology (SAAIP), IJCNLP, Chiang Mai, Thailand, pp. 51–58, 2011.

[10]   Y. Choi, Y. Kim,S-H. Myaeng, "Domain-specific Sentiment Analysis using Contextual Feature Generation," In: Proceedings of the 1st International CIKM Workshop on Topic-Sentiment Analysis for Mass Opinion Measurement(TSA'09), Hong Kong – China, pp. 37-44, 2009.

[11]   S-K. Kim, E. Hovy, "Determining the sentiment of opinions," In: Proceedings of the 20th international conference on computational linguistics (COLING 2004), Geneva, Switzerland. pp. 1367–1373, 2004.

[12]   M. Elhawary,M. Elfeky, "Mining Arabic Business Reviews," In: Proceedings of the2010 IEEE International Conference on Data Mining Workshops; pp. 1108-1113, 2010.

[13]   Pang B, Lee L 2008 Opinion mining and sentiment analysis. Foundations and trends in information retrieval, 2(1-2):1-135.

[14]   Pang B, Lee L, Vaithyanathan S 2002 Thumbs up? Sentiment classification using.

[15]   Machine learning techniques. Proc. ACL-02 Conf. on Empirical methods in natural language processing, 10: 79-8.