Ranking Beauty Clinics in Riyadh using Lexicon-Based Sentiment Analysis and Multiattribute-Utility Theory

Zuhaira M. Zain[1]*, Aya A. Alhajji[2], Norah S. Alotaibi[3], Najwa B. Almutairi[4]
Alaa D. Aldawas[5], Muneerah M. Almazrua[6], Atheer M. Alhammad[7]

Information Systems Department
College of Computer and Information Sciences
Princess Nourah Bint Abdulrahman University
Riyadh, Saudi Arabia

Abstract—In recent years, the amount of beauty-related user-created content has steadily increased. Digital beauty-clinic reviews have major impact on user preferences. In supporting user selection decisions, ranking beauty clinics via online reviews is a valuable study subject, although research on this problem is still fairly limited. Sentiment analysis is a very important subject in the research community to evaluate a predefined sentiment from online texts written in a natural language on a particular topic. Recently, research on sentiment analysis for the Arabic language has become popular, since the language has become the fastest-growing language on the web. However, most sentiment-analysis tools are designed for the Modern Standard Arabic language, which is not widely used on social-media platforms. Moreover, the number of lexicons designed to handle the informal Arabic language is restricted, especially in the beauty-clinic-related domain. Besides, numerous sentiment-analysis studies have concentrated on improving the accuracy of sentiment classifiers. Studies about choosing the right company or product on the basis of the results of sentiment analysis are still missing. In decision-analysis domain, the multiattribute-utility theory has been extensively used in selecting the best option among a set of alternatives. Thus, this research aims to propose a systematic methodology that can develop a beauty-clinic-domain-related sentiment lexicon in Saudi dialect, perform sentiment analysis on online reviews of 10 beauty clinics in Riyadh based on the built lexicon, and feed the lexicon-based sentiment analysis results to the multiattribute-utility theory method to evaluate and rank the beauty clinics. Results showed that the Abdelazim Bassam Clinic is Riyadh's best beauty clinic on the basis of the proposed method. The research not only impacts data analysts regarding how to rate beauty clinics on the basis of lexicon-based sentiment-analysis results, but also directs users toward selecting the best beauty clinic.

Keywords—*Arabic language; beauty clinics; ranking; Lexicon-based; machine learning; sentiment analysis; multiattribute-utility theory*

I. INTRODUCTION

The number of women who undergo cosmetic procedures in Saudi Arabia has grown significantly over the last few years. Liposuction, laser hair removal, Botox, and filler treatments are all classified as cosmetic procedures. Exposure to cosmetic procedures has increased in Saudi Arabia due to media penetration, ever-changing beauty standards, the popularity of the fashion and film industries, per capita income growth, reduced prices for cosmetics, and increasing obesity, thereby generating a high demand for cosmetic treatment in Saudi Arabia [1]. This demonstrates that cosmetic procedures are becoming a trend in Saudi Arabia not only for women, but also for men to some extent; one study showed that 90.5% of clients are female and 9.5% of them are male. Therefore, the number of beauty clinics has also relatively increased [2]. However, according to the Saudi Press Agency [3], many beauty clinics took advantage of this trend in an unpleasant way that resulted in many distortions and legal problems, especially related to unlicensed cosmetic surgeries carried out in beauty salons and unlicensed beauty clinics that used untrusted medical devices. This phenomenon is a growing concern for Saudi women when choosing an appropriate clinic, as there is no clear study or analysis of beauty clinics in Saudi Arabia, especially in Riyadh.

Currently, with the advent and rapid growth of social-media apps, people have begun revealing their views and opinions on, for example, goods, social issues, and policies on the Internet, which has increased the number of user-generated reviews containing rich opinions and sentiment information. Opinion mining, often referred to as sentiment analysis, is a research field that seeks to examine individual perceptions or opinions toward entities such as topics, people, problems, organizations, or events. Sentiment analysis (SA) is a way of classifying text into positive, negative, or neutral sentiments that deal with subjective statements. SA is common because people tend to take advice from other people when making successful investments. It also allows for suppliers to provide insight into the success of their products or services [4].

One of the most commonly utilized techniques is lexicon-based SA. Using this technique, sentiment is classified on the basis of sentiment lexicons. Sentiment lexicons are series of terms correlated with positive- or negative-sentiment orientation [5,6]. They are also called polar or opinion terms. There were several studies performed that built sentiment lexicons for the Modern Standard Arabic language [7–10]. However, they are not relevant to views or comments extracted from social-media platforms, which mostly use the

*Corresponding Author

informal Arabic language. The informal Arabic language ignores standard rules of spelling and grammar. Furthermore, it can vary from one region to another because each region has its own dialect. Several studies developed informal Arabic lexicons in the Lebanese [11], Algerian [12], and Saudi dialects [13,14], but they might not be relevant to the beauty-clinic-related context. In addition, most sentiment-analysis studies focused on improving accuracy. There is still a lack of studies providing information on how to select the best product or organization on the basis of sentiment-analysis results.

There is an established method called multicriteria decision making (MCDM) that can help decision-makers to make decision on a set of alternatives. Some well-known MCDM techniques include multiattribute-utility theory (MAUT) [15], VlseKriterijumska Optimizcija I Kaompromisno Resenje (VIKOR) [16], analytic hierarchy process (AHP) [17], and Technique for Order Preference by Similarity to the Ideal Solution (TOPSIS) [18]. Research on MCDM has shown that it is challenging to determine which technique is better because no technique can be the best among the existing techniques [19,20]. However, because of its simplicity, MAUT was commonly used [21–23]. MAUT is a well-established form of decision analysis that explicitly discusses how an alternative can be chosen from a collection of alternatives. This approach allows for the decision maker to separately assess alternatives for each attribute. The decision maker allocates relative weights to different attributes. Then, a formal model is used to combine and aggregate values and weights to produce a global assessment of each alternative [15]. A few examples of successful MAUT applications include its recent use to rank a number of bridges [24], select iron and steel suppliers [25], and identify attributes for inclusion in the framework of an education recommendation system [26]. Nevertheless, to the best of our knowledge, it has not been used to analyze decisions based on sentiment analysis results.

To tackle the aforementioned drawbacks, this study proposes a research method that can build a beauty-clinic-domain-related sentiment lexicon in Saudi dialect, conduct sentiment analysis based on the developed lexicon and pass the results to the multiattribute-utility theory method to assess and rank the beauty clinics. This study facilitates potential clients in selecting an experienced and reputable beauty clinic based on reviews from other people with previous experience in cosmetic procedures by extracting their opinions from Twitter, which is one of the most famous social-media platforms where people express their views. The key contributions of this study are as follows:

- A new sentiment lexicon was built for the informal Arabic language (Saudi dialect) from the Twitter accounts of 10 beauty clinics in Riyadh using a manual approach to address the irrelevant sentiment-lexicon problem;

- Online reviews from the Twitter accounts of 10 beauty clinics were preprocessed by performing tokenization,

normalization, cleaning, stemming, and stop-words removal; and classified into sentiment using the newly built sentiment-lexicon;

- The performance of the 10 beauty clinics was evaluated by applying multiattribute-utility theory using the results of lexicon-based sentiment analysis as inputs to show their ranking.

The remainder of this study is structured as follows: Section 2 outlines previous work related to the study; Section 3 elaborates on the methodology used for the research; Section 4 introduces the findings of the study; Section 5 addresses the sensitivity analyses; lastly, Section 6 provides the study conclusions.

II. RELATED WORK

A. Sentiment Lexicon

A sentiment lexicon is one of the most useful means for carrying out sentiment-analysis research for any language. It is an essential tool for both unsupervised (lexicon-based) and supervised (machine-learning) classifiers. It is used by many researchers to generate unsupervised sentiment models or training features in supervised approaches to train machine-learning algorithms. A sentiment lexicon is a set of opinion words or polar terms correlated on the basis of their orientation, i.e., positive, negative, or neutral. The sources and approaches applied to non-English sentiment-lexicon development are graphically illustrated in Fig. 1. For further information on each method, readers may refer to the work of Kaity and Balakrishnan [27].

Al-Twairesh et al. [28] used a translation-based approach to generate comprehensive Arabic sentiment lexicons (AraSenti-Trans) employing the MADAMIRA tool [29], which recognizes Arabic terms in tweets and eliminates tweets containing non-Arabic terms and dialects. They applied two lexicons constructed by Hu and Liu [30], and Wilson et al. [31] as sentiment-orientation tools after preprocessing. They used MADAMIRA's English glossary to determine word polarity upon comparing both lexicons in [30] and [31]. They also applied a frequency-based approach to build the AraSenti-PMI lexicon by adopting the pointwise-mutual-information (PMI) measure to differentiate words as either positive or negative in a corpus. Mahyoub et al. [32] used a relationship-based approach by introducing an algorithm that assigns sentiment scores to Arabic WordNet entries for Arabic sentiment-lexicon construction. Using synset relationships, a semisupervised-learning algorithm was applied to enhance the number of terms in Arabic WordNet. Badaro et al. [7] used a merger-based approach to combine four existing sentiment lexicons (Standard Arabic Morphological Analyzer [33], English WordNet, English SentiWordNet [34,35], and Arabic WordNet [36]) into a new Arabic sentiment lexicon (ArSenL). However, these lexicons were built for standard Arabic, which is not relevant to the language used on social media, i.e., informal Arabic.
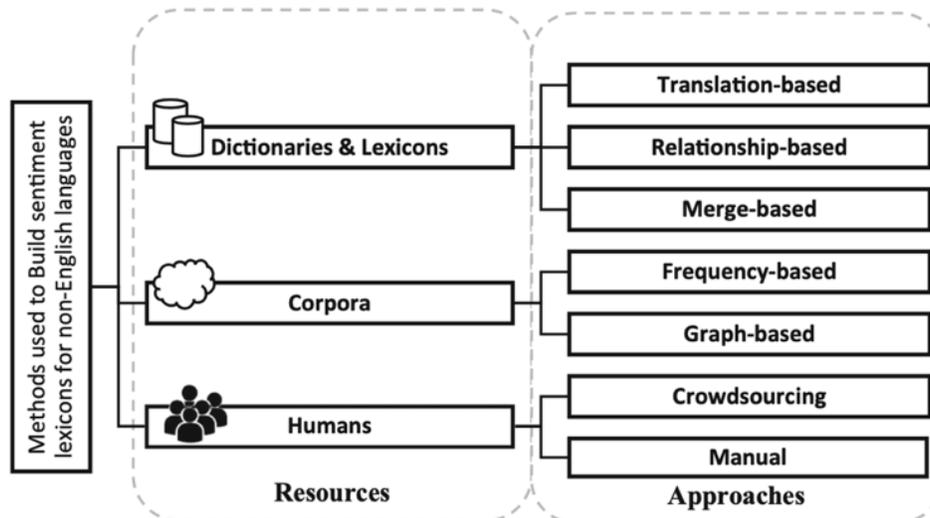
Fig. 1.   Resources and Approaches applied to non-English Sentiment-Lexicon Development [27].

Duwairi et al. [37] adopted a crowdsourcing approach to build three sentiment lexicons: a lexicon mapping the Jordanian dialect to Modern Standard Arabic (MSA), a lexicon mapping Arabizi terms to MSA, and an emoticon lexicon. Abdul-Mageed et al. [38] applied a manual approach to develop a 3982 adjective sentiment lexicon as part of the SAMAR system built in MSA and Arabic dialects to examine Arabic subjectivity and sentiments. Al-Ghaith [14] built the SaudiSentiPlus sentiment lexicon containing 7139 Saudi dialect terms by first using an automatic translation of English sentiment lexicons previously generated by [30] and [39]. He then manually extracted all Saudi dialect terms from Twitter sentiment data. However, the lexicons may not be applicable to the beauty-clinic-related context. Despite it being considered time-consuming and costly, many researchers still use a manual approach, particularly those studying sentiment analysis in languages lacking lexical resources.

B. Sentiment Analysis

Sentiment analysis is computational analysis of sentiments, opinions, and emotions in a text related to a specific subject [6]. It assists in attaining many objectives, such as observing public sentiments toward political movements, measuring customer satisfaction [40], establishing market intelligence [41], and predicting sales for a product. Therefore, researchers took advantage of SA and developed it for various purposes in different fields and languages [42,43]. There are four sentiment-analysis tasks: sentiment classification, subjective classification, opinion-spam detection, and review usefulness. SA classification can be done on three levels: the document level, which attempts to show the polarity of an entire document opinion focusing on a single topic (e.g., product or place) from a single opinion holder; the sentence level, which assumes the sentence is a distinct unit that comprises only an opinion; and the feature level, also called the entity level, which seeks to assign polarity to multiple reviews using extracted features [44–48]. These processes can be completed by applying three main techniques: a lexicon-based, machine-learning, or hybrid technique [49]. The lexicon-based technique counts and weighs sentiment-related words to enable the adoption of a lexicon to perform sentiment analysis. The machine-learning (ML) technique predicts the polarity of a large number of posts using training datasets or annotated corpora. The hybrid technique combines the lexicon-based and machine-learning techniques, whereby the lexicon-based classifier builds a dataset by labeling each tweet on the basis of a sentiment lexicon, and the machine-learning-based classifier is trained and tested using the dataset generated from the lexicon-based approach [49,50].

1) Lexicon-Based approach: A lexicon-based approach, known as the dual-polarity algorithm, was developed by El-Beltagy and Ali [51] to determine the weights of Egyptian dialect sentiment lexicon terms, achieving 70% accuracy. Abdulla et al. [52] constructed a lexicon-based sentiment-analysis method by considering negation and intensification to measure text polarity. They applied a basic lexicon-based approach on an Arabic corpus collected from Twitter and Yahoo!-Maktoob by calculating positive and negative terms in a given text, and then they categorized the sentences using a larger category. Applying various lexicon scalability stages, they achieved 70.05% accuracy. Ayyoub et al. [8] developed a sentiment lexicon of approximately 120,000 Arabic terms extracted from Twitter, and developed a lexicon-based sentiment-analysis method using a predicate calculus. Their proposed method demonstrated better predictive accuracy (86.89%) compared to that of the keyword-based method. Assiri et al. [10] built a Saudi dialect lexicon manually that includes 14,000 sentiment words. They then built a weighted lexicon-based classifier. Their proposed classifier removed associations between polarity and nonpolarity terms for a dataset, and then weighted these terms by using their associations. They applied new rules for processing certain linguistic features such as negation and supplication. Their enhanced lexicon-based sentiment-analysis classifier achieved 81% accuracy.

2) Machine-Learning approach: Hodeghatta [53] conducted sentiment analysis using the naïve Bayes and MaxEnt algorithms to categorize Twitter messages related to Hollywood movies as positive, negative, and cognitive sentiments across various regions of various countries. The classifiers were tested for both unigrams and bigrams. The MaxEnt classifier with unigrams achieved the highest accuracy of 84%. Jokhio et al. [54] determined the sentiment polarity (positive, negative, or neutral) of English tweets related to plastic-surgery treatments in the United States using a naïve Bayes algorithm. The experimental results demonstrate the importance of the suggested method, which may allow potential groups of people who want to have plastic-surgery treatments to make better decisions. Recently, Naseem et al. [55–57] proposed Deep Intelligent Contextual Embedding (DICE) [55], Hybrid Words Representation [56] and Transformer-based Deep Intelligent Contextual Embedding (DICET) [57] models to improve tweet quality by considering polysemy, syntax, semantics, and out of vocabulary (OOV) words; and handling the noise within the textual context. The evaluation of the models executed on airline-related datasets indicated that their proposed models improved the accuracy for tweets classification with an average accuracy of 93.5%, 94.2%, and 94.6%, respectively. The results verified that their models can successfully handle low-quality data and language complexities. However, these studies only considered tweets written in the English language.

3) Hybrid approach (Lexicon-Based + ML): Yoo and Nam [58] proposed a hybrid approach to Korean sentiment analysis, where a lexicon-based classifier was used to locally parse sentiment components to identify opinions, while a naïve Bayes classifier was used to categorize the text in a dictionary by training the MUSE (Multilingual Sentiment Lexica and Sentiment-Annotated Corpora) opinion corpus. Aldayel and Azmi [59] combined a lexicon-based technique and a machine-learning-based algorithm to detect the polarities of Arabic tweets. Through this approach, the lexical-based classifier was used to label the training data from a manually built sentiment lexicon. The output of the lexicon-based classifier was adopted as training data for the support-vector-machine (SVM) classifier.

However, all aforementioned studies [8,10,51–59] solely focused on how to improve the accuracy or performance of the sentiment classifier, which is not the main focus in this study. There is still a lack of studies providing information on how to select the best product or organization on the basis of sentiment-analysis results.

C. Multiattribute-Utility Theory

Multiattribute-utility theory (MAUT) is a well-established decision-analysis approach that discusses precisely how to choose an alternative from a set of alternatives. Although the functional implementation of the multiattribute-utility technique can differ, the process involves the following steps [15]:

1) Identifying alternatives and value-relevant attributes: The first process is to define the alternatives available and their most important attributes.

2) Separately assessing each alternative in terms of each attribute: Each alternative is assessed in terms of each attribute set out in Step 1.

3) Allocating relative weights to attributes: The weights for each attribute are allocated using a direct-rating method. The total rating for each attribute is calculated and transformed into weights using the following formula:

$$w_i = \frac{w'_i}{\sum_{i=1}^{n} w'_i},\tag{1}$$

where w_i is the weight for each attribute, and w'_i is the relative importance rating score for each attribute. Thus, individual weights are obtained that sum to 1 for each attribute, as is conventional in the MAUT [15].

4) Aggregating attribute weights and single-attribute assessments of alternatives for a global assessment of alternatives: Weights and single-attribute values or functions are aggregated by adopting the weighted linear additive preference function [15] shown in (2).

$$v(A) = \sum_{i=1}^{n} w_i \, v_i(A_i),\tag{2}$$

where $v_i(A_i)$ is the value of alternative A in terms of the i-th attribute, w_i is the importance weight of the i-th attribute, and n is the number of different attributes [15].

5) Conducting sensitivity analyses and offering recommendations: In the previous step, multiattribute utilities are determined for all alternatives. The best alternative should be that with the highest multiattribute utility. Sensitivity analyses are carried out in the last stage of the process to determine the stability of the results. The effect of various values and weights on the multiattribute utility of the available alternatives can then be calculated. Different values and weights can be obtained using various methods of elicitation. For instance, both the direct-rating method and the bisection approach may be applied to obtain values. Thus, multiattribute utilities are measured twice: once using the values calculated using the direct-rating method, and then using the values calculated using the bisecting tool. Then, the resulting multiattribute utility for the available alternatives can be evaluated, and the stability of the results can be calculated. An alternative way of understanding the robustness of the results is to apply a different weighting method. For instance, the equal-weight method can be adopted. This method simplifies the selection process by avoiding information on the relative importance of each attribute [60,61]. This means that the approach assumes that all attributes are of equal weight. For instance, for four attributes, the weight of each attribute would be 1/4 = 0.25. The total value for each alternative is attained by summing the values in terms of each attribute. The rank-reciprocal rule and rank-sum weighting techniques can also be used to assign weights [15].

III. METHODS

The research methodology of this study, as depicted in Fig. 2, comprises six phases: (1) collecting data, (2) constructing sentiment lexicon, (3) preprocessing, (4) classifying tweets, (5) developing classification models, and (6) ranking beauty clinics. The details of each phase are outlined in the subsections.

A. Phase 1: Collecting Data

In this phase, data were collected using the Twitter application programming interface (API), which is an interface for researchers to collect data from a given social-media service. Opinions were aggregated using the *Tweepy* library by searching for tweets related to the 10 selected beauty clinics (Panorama, Pearl, Alzallal, Obagi, Abas, Derma, Kesaaee, AbdelAzim Bassam, Adma, and The Clinic). These beauty clinics are the most popular in Riyadh. The overall number of aggregated tweets was 780.

B. Phase 2: Constructing Sentiment Lexicon

A sentiment lexicon is a table where each word is associated with its polarity, which indicates each word's orientation (positive, neutral, or negative). In this phase, an informal Arabic sentiment lexicon in the Saudi dialect was built by manually extracting sentiment words from the collected dataset. Each sentiment word was assigned its polarity. Afterward, the lexicon was used to assign polarity to each tweet using its score.

As this project was conducted in the informal Arabic language, many tweets contained emoji. Since most emoji change the meaning of sentences, an emoji lexicon was also built. A function from the Python library called *extract_emojis()* was used to find the emoji in our dataset.

In order to simplify the classification process in Phase 4, the lexicon was separated into four groups: positive- and negative-word lexicons, and positive- and negative-emoji lexicons. Each of these lexicons was stored in a different list.

C. Phase 3: Preprocessing Data

In this phase, the dataset was preprocessed using the following techniques:

- Tokenization: Data were transformed from the sentence level to the word level by dividing the text into a set of meaningful tokens using the *codecs* and *nltk* libraries, and the *word_tokenize()* function.

- Text cleaning: All non-Arabic words, special Twitter characters, and usernames (for example, # and @), and irrelevant information (for example, universal resource locators (URLs)) were removed. In addition, retweets were eliminated.

- Text normalization and eliminating repeated letters: In this step, some Arabic letters were transformed into general letters, such as unifying the letter *alif* by replacing إ, أ, and آ with ا. In addition, repeated letters such as مرررة and مرة were removed. This process was done using the *re*, *sys*, *string*, and *argparse* libraries.

- Stop-word removal: Words with no meaning or that do not hold information and have no effect on the output, such as conjunctions, articles, and prepositions, including في/fi (in), من/min (of), and على/ala (on), were removed. The stop-word list was imported from the *nltk* corpus, which contains MSA stop words. Forty-nine new informal Arabic stop words were added to the list.

- Text stemming: Numerous Arabic words have a similar stem. For example, يكره/yakrah (he hates), تكره/takrah (she hates), اكره/akrah (I hate), يكرهون/yakrahon (they hate) all originate from كره/kurh (hate); thus, all words were replaced with their كره/kurh (hate) stem using the *nltk ISRIStemmer* to eliminate redundant terms.
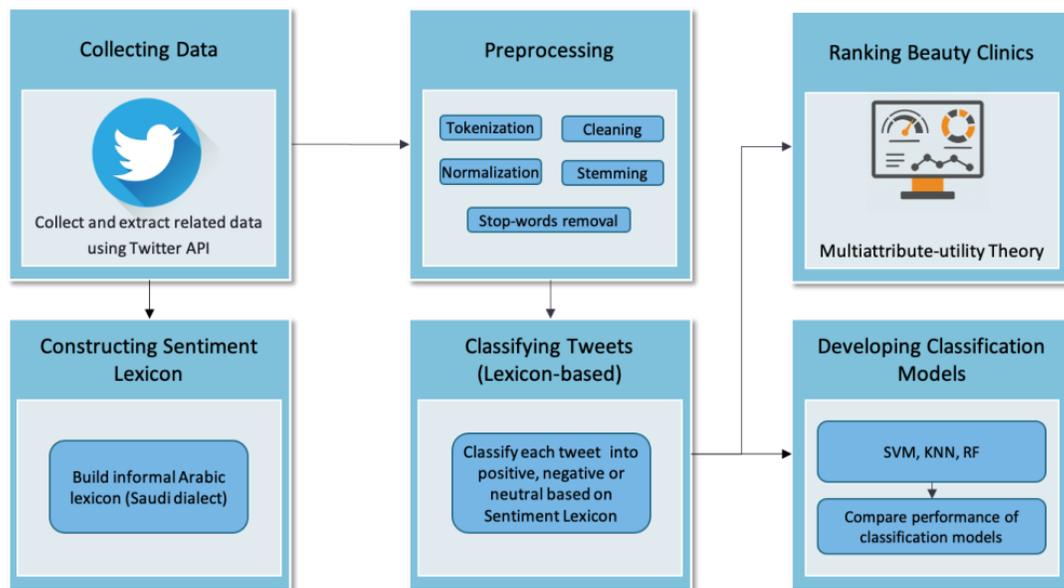

Fig. 2. Research Methodology.

D. Phase 4: Classifying Tweets

Algorithm 1 [59] was applied to classify the preprocessed unannotated tweets as positive, negative, or neutral. The algorithm takes the dataset containing tweets and tokenizes it; then, it checks each token's score from the built lexicon and adds it for each sentence. If the overall score for the tweet is greater than zero, a positive label is assigned; otherwise, a negative label is assigned.

Algorithm 1

Input: preprocessed unannotated tweets

Output: set of classified and annotated tweets along with their sentiment (positive or negative), and a set of unclassified tweets (corresponding to neutral sentiment).

Begin

For each tweet t in training set do

W = tokenize tweet t into a list of words $w_1, w_2, \ldots$ and

emoji $e_1, e_2, \ldots$

score = 0

For each w_i, e_i in W do

Stem w_i, e_i

If (w_i is in Lexicon) score = score + positivity_degree(w_i)

Else If (e_i is in Lexicon) score = score + positivity_degree(e_i)

End for

If (score > 0) return POSITIVE

Else return score < 0 ? NEGATIVE:NEUTRAL

End

In order to refer to our four lexicon groups, a function called *tweets_polarity* was created to determine the polarity of each tweet. This function takes one parameter (i.e., a tweet) and returns its polarity. At first, each tweet is tokenized and stored into a tokenized-tweet list, W. Second, the *score* variable is set to zero. Third, a *for* loop is created to run through each word and emoji in W. The body of the *for* loop contains an *if* condition to check if the word belongs to a certain lexicon; for example, if the word belongs to the positive-word lexicon, the *score* variable is increased by one. Fourth, another *if* condition enables determining the polarity value using the *score* variable; for example, if the *score* variable was greater than zero, the polarity was set to positive.

A *for* loop was created to run through each entry in the tweet list, followed by storing the value of that entry as a variable and passing it along to the *tweets_polarity* function before calling it. This function is supposed to return the polarity value for a tweet each time it is called. Lastly, the tweets and their polarity values were saved to a dataset. This dataset was used in Phase 5 to train and test three machine-learning classifiers. It was also used in Phase 6 to calculate the total score and rank the beauty clinics.

E. Phase 5: Developing Classification Model

In this phase, classification models were built on the basis of three popular classifiers: random forest (RF), support vector machine (SVM), and K-nearest neighbor (KNN). The models were trained and tested using Python. The dataset generated from Phase 4 was divided into training and testing sets in a 70:30 ratio. All three classifiers were set to their default values. Next, the performance of the models was compared on the basis of their accuracy.

F. Phase 6: Ranking Beauty Clinics

In this phase, the multiattribute-utility theory technique was applied in the following steps:

1) Identifying alternatives and value-relevant attributes: In this study, we evaluated the 10 most popular beauty clinics in Riyadh on the basis of three attributes generated from lexicon-based sentiment analysis: positive, negative, and neutral tweets.

2) Separately assessing each alternative in terms of each attribute: The 10 beauty clinics were evaluated using the three attributes generated from the lexicon-based sentiment analysis as a function of the total count of positive, negative, and neutral tweets.

3) Allocating relative weights to the attributes: The weights for each attribute were assigned using the direct-rating method. An online questionnaire was designed where the decision makers (beauty-clinic users) were invited to rate the relative importance of the three attributes on a five-point Likert scale, where 1 denotes the least and 5 denotes the most important attribute. The total rating for each attribute was computed and transformed into weights using (1).

4) Aggregating the weights of attributes and the single-attribute assessments of alternatives to calculate the global assessment of alternatives: Weights and single-attribute values were aggregated using (2).

5) Running sensitivity analyses and generating recommendations: The alternatives were ranked on the basis of the aggregation score obtained in Step 4. The alternative with the highest aggregation score was assigned the highest rank; the alternative with the lowest score was assigned the lowest rank. The alternative with the highest score was considered the most preferred. Sensitivity analyses were run to evaluate the stability of the results. In this study, two sensitivity analyses were conducted: (1) using only the single attribute with the highest weight, and (2) using equal weight. In the first, alternatives were ranked on the basis of the rating score for the attribute with the highest weight. In the second, since we used the equal-weight technique, the weight for each attribute was 0.33. The global value for each alternative was computed, and the alternatives were ranked on the basis of their global score. The ranking results using these techniques were compared with the result using the proposed technique.

In order to direct us in the presentation of the results of our analysis, the six following research questions were raised:

- RQ1: What is the structure of the built sentiment lexicon?

- RQ2: What is the total number of positive, negative, and neutral tweets for each beauty clinic?

- RQ3: Which ML classifiers perform best?

- RQ4: Which clinic is the most preferred by clients?

- RQ5: What is the impact of using the single-attribute utility?

- RQ6: What is the impact of using different weighting methods?

IV. RESULTS

A. RQ1 Answer: Structure of Sentiment Lexicon

The sentiment lexicon generated from the 780 tweets collected in this study consisted of 658 sentiment words: 276 negative, 276 positive, and 106 neutral words. Table 1 shows a sample of the informal Arabic word lexicon.

It also consisted of 150 sentiment emoji: 72 positive, 54 negative, and 24 neutral emoji. Table 2 presents an example of the emoji lexicon. Since neutral words and emoji were assigned a polarity of zero, they were not inserted into the lexicon list.

B. RQ2 Answer: Total Number of Positive, Negative, and Neutral Tweets for each Beauty Clinic

We were unable to present the results of tweet classification because we are bound to Twitter's privacy policy that prevents the publication of the original tweets that were collected. Table 3 portrays the number of tweets for each beauty clinic that were classified as positive, negative, or neutral. The total count of positive, negative, and neutral tweets was considered as the rating score for each clinic in terms of each attribute.

C. RQ3 Answer: Best ML Classifier

Table 4 demonstrates that SVM outperformed RF and KNN in classifying the tweets as positive, negative, or neutral sentiments with 72.1% accuracy. This result was expected, and it was deemed acceptable for the small size of the training dataset. It may have achieved better performance if we had increased the size of the training dataset.

D. RQ4 Answer: Most Preferred Beauty Clinic

Table 5 shows the total relative importance of each attribute as rated by 46 beauty-clinic users. The weights for each attribute were calculated by using (1). Positive tweets had the highest weight value (0.5), followed by neutral tweets (0.4) and negative tweets (0.1). These weight values were used to aggregate the rating score for each beauty clinic.

Table 6 demonstrates the total rating score for each beauty clinic that was computed using (2). The global scores for each beauty clinic were sorted in descending order, and a bar graph was plotted to show the rank.

TABLE I. SAMPLE OF INFORMAL ARABIC WORD LEXICON

Word	Polarity	Word	Polarity
تهبل/t'habel (amazing/female form)	1	مقرف/mogref (disgusting)	−1
يهبل/yhabel (amazing/male form)	1	زفت/zift (bad)	−1
يخقق/ykhageg (outstanding)	1	زحمه/zahma (crowded)	−1
رهيب/raheeb (awesome)	1	بعيد/baeed (far)	−1
احسن/ahsan (better/best)	1	سيئ/saye' (bad)	−1
افضل/afdhal (better/best)	1	سيء/saye' (bad)	−1

TABLE II. SAMPLE OF EMOJI LEXICON

Emoji	Polarity	Emoji	Polarity
👍	1	☹	−1
♡	1	😫	−1
😂	1	💔	−1
👌	1	👎	−1
❤	1	😣	−1
👍	1	😟	−1

TABLE III. TWEET CLASSIFICATION

Clinic	Positive Tweets, x_1	Negative Tweets, x_2	Neutral Tweets, x_3	Total Tweets
Panorama Clinic, A_1	8	45	43	96
Pearl Clinic, A_2	3	12	19	34
Alzallal Clinic, A_3	10	18	35	63
Obagi Clinic, A_4	6	6	22	34
Abas Clinic, A_5	23	18	77	118
Derma Clinic, A_6	36	28	98	162
Kesaaee Clinic, A_7	20	7	26	53
AbdelAzim Bassam, A_8	63	18	78	159
Adma Clinic, A_9	4	3	2	9
The Clinic, A_{10}	26	4	22	52
Total, $v_i(A)$	**199**	**159**	**422**	**780**

TABLE IV. ACCURACY OF MACHINE-LEARNING CLASSIFIERS

Classifier	Accuracy (%)
Random forest (RF)	67.2
Support vector machine (SVM)	72.1
K-nearest neighbor (KNN)	67.2

TABLE V. WEIGHTS FOR EACH ATTRIBUTE

	Positive Tweets	Negative Tweets	Neutral Tweets	Total
Total relative importance rating score	226	46	180	452
w_i	226/452 = 0.5	46/452 = 0.1	180/452 = 0.4	1

TABLE VI. TOTAL AGGREGATION SCORE FOR BEAUTY CLINICS

Clinic	$w_1v_1(A_i)$	$w_2v_2(A_i)$	$w_3v_3(A_i)$	$v(A_i)$
Panorama Clinic, A_1	4	4.5	17.2	25.7
Pearl Clinic, A_2	1.5	1.2	7.6	10.3
Alzallal Clinic, A_3	5	1.8	14	20.8
Obagi Clinic, A_4	3	0.6	8.8	12.4
Abas Clinic, A_5	11.5	1.8	30.8	44.1
Derma Clinic, A_6	18	2.8	39.2	60.0
Kesaaee Clinic, A_7	10	0.7	10.4	21.1
AbdelAzim Bassam, A_8	31.5	1.8	31.2	64.5
Adma Clinic, A_9	2	0.3	0.8	3.1
The Clinic, A_{10}	13	0.4	8.8	22.2

Fig. 3 illustrates the ranking of the 10 Riyadh beauty clinics. AbdelAzim Bassam Clinic was found to be the most preferred among all beauty clinics investigated in this study on the basis of its rank.
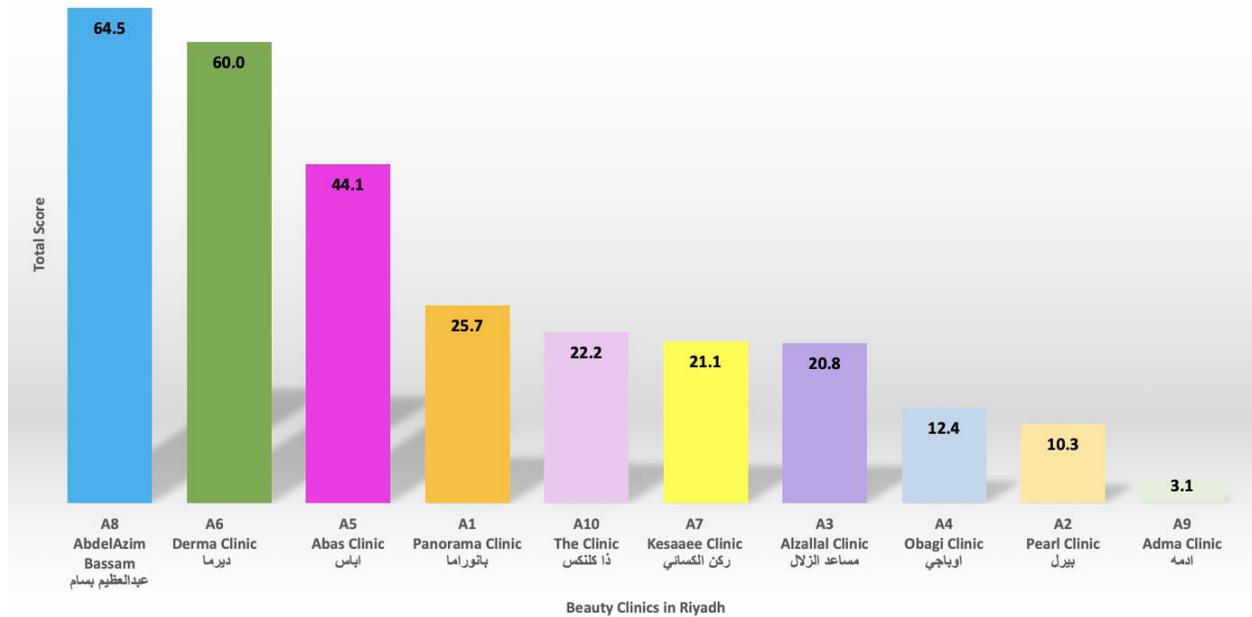
Fig. 3. Ranking of Riyadh's Beauty Clinics

V. DISCUSSION

A. RQ5 Answer: Impact of using the Single-Attribute Utility

Table 7 presents the results of the first sensitivity analysis that was conducted using a single attribute. Results showed that the existing ranking of the alternatives changed for five pairs of alternatives: A_1–A_3, A_1–A_7, A_1–A_9, A_2–A_9, and A_5–A_{10}. Despite these changes in the ranking of some pairs of alternatives, the first-ranked alternative remained unaffected.

B. RQ6 Answer: Impact of using Different Weighting Methods

Table 8 displays the results of the second sensitivity analysis that was conducted using equal weight. Results showed that the existing ranking of the alternatives changed for five pairs of alternatives: A_2–A_4, A_3–A_7, A_3–A_{10}, A_6–A_8, and A_7–A_{10}. The impact of using equal weight was more critical because it affected the first-ranked alternative.

TABLE VII. SENSITIVITY ANALYSIS USING A SINGLE ATTRIBUTE

Method	Ranking
Proposed (multiattribute utility)	$A_8 \succ A_6 \succ A_5 \succ A_1 \succ A_{10} \succ A_7 \succ A_3 \succ A_4 \succ A_2 \succ A_9$
Single-attribute utility (positive tweets)	$A_8 \succ A_6 \succ A_{10} \succ A_5 \succ A_7 \succ A_3 \succ A_1 \succ A_4 \succ A_9 \succ A_2$

TABLE VIII. SENSITIVITY ANALYSIS USING EQUAL WEIGHT

Method	Ranking
Proposed (relative importance)	$A_8 \succ A_6 \succ A_5 \succ A_1 \succ A_{10} \succ A_7 \succ A_3 \succ A_4 \succ A_2 \succ A_9$
Equal weight (0.33)	$A_6 \succ A_8 \succ A_5 \succ A_1 \succ A_3 \succ A_7 \succ A_{10} \succ A_2 \succ A_4 \succ A_9$

C. Validity Threats

There were three threats to validity considered in this study: threats to construct, internal, and external validity.

Threats to construct validity were related to the performance measure used in this study. Machine-learning performance was measured on the basis of accuracy. Many other performance metrics can be used, such as F-measure, sensitivity, and precision. However, accuracy is widely used by many researchers when measuring the performance of lexicon-based or machine-learning classifiers in sentiment-analysis research.

Threats to internal validity were related to uncontrolled internal factors that might have influenced the experiment results. These internal factors would have occurred during method implementation. In order to minimize these threats, we first used a standard lexicon-based algorithm proposed by [59] to build the training dataset. Second, we chose three well-known classifiers provided by scikit-learn libraries [62] (RF, SVM, and KNN) to train and test the tweet dataset. We applied default settings for all three classifiers. Third, the formula used for MAUT was taken from an established resource [15].

Threats to external validity were related to the possibility of generalizing the results of this study. One possible factor was the use of a dataset from 10 beauty clinics in Riyadh. This dataset was very limited and domain-specific. However, this study focused on evaluating 10 beauty clinics in Riyadh; thus, the best way to reduce this threat was to select the Twitter accounts of the 10 most popular beauty clinics in Riyadh. Another possible factor constituted the methods used for sensitivity analysis. In this study, we ran sensitivity analyses using the single-attribute evaluation technique and the equal-weight method. There are various other methods that can be used, such as VIKOR, AHP, TOPSIS, and the rank-sum

weighting method. However, the methods used in this analysis are among numerous suggested in the literature, and they are commonly used because of their simplicity.

VI. CONCLUSION

In this study, a novel method for ranking beauty clinics in Riyadh was proposed. It started with data collection from 10 beauty clinics in Riyadh using the Twitter API. The dataset was manually transformed into a sentiment lexicon. Then, data were preprocessed by applying tokenization, cleaning, normalization, stop-word removal, and stemming techniques. Next, preprocessed tweets were classified as positive, negative, or neutral using a lexicon-based approach. Lastly, the rating score for each beauty clinic was computed using multiattribute-utility theory to rank the beauty clinics.

Results showed that Abdelazim Bassam Clinic is Riyadh's best beauty clinic on the basis of the proposed method. This study will impact clients when choosing the best beauty clinic. Moreover, it can assist beauty-clinic owners in understanding how they are faring with their clients, as it gives them a better picture of how they stack up against their competitors, thereby providing them with an opportunity to improve their services. Furthermore, this study provides data analysts with an example of how to rate beauty clinics by using lexicon-based sentiment-analysis results.

In the future, we will increase the number of words and emoji in our sentiment lexicon. Furthermore, we will combine some machine-learning algorithms and evaluate their performance in classifying tweets as a function of their polarity. An aspect-based sentiment analysis will be applied to obtain additional attributes to evaluate the performance of beauty clinics.

REFERENCES

[1] Technavio Research, "Global Cosmetic Surgery Market 2019|2023| Evolving Opportunities with ALLERGAN and Bausch Health| Technavio," 2019.

[2] S. E. Alharethy, "Trends and demographic characteristics of Saudi cosmetic surgery patients," Saudi Medical Journal. vol. 38, no. 7, pp. 738–741, 2017, doi: 10.15537/smj.2017.7.18528.

[3] Saudi Press, "Medicine and health prohibit a medical device used in a beauty clinic," 2017.

[4] M. T. Khan, M. Durrani, A. Ali, I. Inayat, S. Khalid, and K. H. Khan, "Sentiment analysis and the complex natural language," Complex Adaptive Systems Modeling. vol. 4, no. 1. Springer, Dec. 01, 2016, doi: 10.1186/s40294-016-0016-9.

[5] S. Ahire, "A survey of sentiment lexicons," 2015, [Online]. Available: https://www.semanticscholar.org/paper/A-Survey-of-Sentiment-Lexicons-Ahire/2522de6022acf2bc7d5c12a9467d4c41f6358920.

[6] W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications: a survey," Ain Shams Engineering Journal. vol. 5, no. 4, pp. 1093–1113, Dec. 2014, doi: 10.1016/j.asej.2014.04.011.

[7] G. Badaro, R. Baly, H. Hajj, N. Habash, and W. El-Hajj, "A large scale Arabic sentiment lexicon for Arabic opinion mining," in Proceedings of the EMNLP 2014 Workshop on Arabic Natural Langauge Processing (ANLP), 2014, pp. 165–173.

[8] M. al Ayyoub, S. B. Essa, and I. Alsmadi, "Lexicon-based sentiment analysis of Arabic tweets," International Journal of Social Network Mining. vol. 2, no. 2, p. 101, 2015, doi: 10.1504/IJSNM.2015.072280.

[9] R. Eskander and O. Rambow, "SLSA: A sentiment lexicon for standard Arabic," in Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, 2015, pp. 2545–2550.

[10] A. Assiri, A. Emam, and H. Al-Dossari, "Towards enhancement of a lexicon-based approach for Saudi dialect sentiment analysis," Journal of Information Science. vol. 44, no. 2, pp. 184–202, Apr. 2018, doi: 10.1177/0165551516688143.

[11] T. Tobaili, M. Fernandez, H. Alani, S. Sharafeddine, H. Hajj, and G. Glavaš, "Senzi: A sentiment analysis lexicon for the latinised Arabic (Arabizi)," in International Conference Recent Advances in Natural Language Processing, RANLP, 2019, vol. 2019-September, pp. 1203–1211, doi: 10.26615/978-954-452-056-4_138.

[12] M. H. Mataoui, O. Zelmati, and M. Boumechache, "A proposed lexicon-based sentiment analysis approach for the vernacular Algerian Arabic," Res. Comput. Sci. vol. 110, 2016.

[13] S. M. Mohammad, M. Salameh, and S. Kiritchenko, "Sentiment lexicons for Arabic social media," in 10th Annual International Conference obn Language Resources and Evaluation, 2016, pp. 33–37, [Online]. Available: http://www.saifmohammad.com/WebPages/ArabicSA.html.

[14] W. Al-Ghaith, "Developing lexicon-based algorithms and sentiment lexicon for sentiment analysis of Saudi dialect tweets," International Journal of Advanced Computer Science and Applications. vol. 10, no. 11, 2019, doi: 10.14569/IJACSA.2019.0101112.

[15] D. von Winterfeldt and W. Edwards, Decision analysis and behavioral research. Cambridge: Cambridge University Press, 1986.

[16] S. Opricovic and G.-H. Tzeng, "Multicriteria planning of post-earthquake sustainable reconstruction," Computer-Aided Civil and Infrastructure Engineering. vol. 17, pp. 211–220, 2002.

[17] T. L. Saaty, "How to make a decision: the analytic hierarchy process," European Journal of Operational Research. vol. 48, pp. 9–26, 1990.

[18] C.-L. Hwang, Y.-J. Lai~, and T.-Y. Liu~, "A new approach for multiple objective decision making," Computers Ops Res. vol. 20, no. 8, pp. 889–899, 1993.

[19] M. Qaradaghi and J. P. Deason, "Analysis of MCDM methods output coherence in oil and gas portfolio prioritization," Journal of Petroleum Exploration and Production Technology. vol. 8, no. 2, pp. 617–640, Jun. 2018, doi: 10.1007/s13202-017-0344-0.

[20] I. Siksnelyte-Butkiene, E. K. Zavadskas, and D. Streimikiene, "Multi-criteria decision-making (MCDM) for the assessment of renewable energy technologies in a household: A review," Energies. vol. 13, no. 5. MDPI AG, Mar. 01, 2020, doi: 10.3390/en13051164.

[21] M. Velasquez and P. T. Hester, "An analysis of multi-criteria decision making methods," 2013.

[22] M. Cinelli, S. R. Coles, and K. Kirwan, "Analysis of the potentials of multi criteria decision analysis methods to conduct sustainability assessment," Ecological Indicators. vol. 46. Elsevier, pp. 138–148, 2014, doi: 10.1016/j.ecolind.2014.06.011.

[23] H. Koornneef, W. J. C. Verhagen, and R. Curran, "A multi-criteria decision making framework for aircraft dispatch assessment," in Advances in Transdisciplinary Engineering, 2017, vol. 5, pp. 11–20, doi: 10.3233/978-1-61499-779-5-11.

[24] Z. A. Bukhsh, I. Stipanovic, and A. G. Doree, "Multi-year maintenance planning framework using multi-attribute utility theory and genetic algorithms," European Transport Research Review. vol. 12, no. 3, 2020, doi: https://doi.org/10.1007/s41109-019-0237-x.

[25] Z. Ma, J. Zhu, and S. Zhang, "Probabilistic-based expressions in behavioral multi-attribute decision making considering pre-evaluation," Fuzzy Optimization and Decision Making. 2020, doi: 10.1007/s10700-020-09335-8.

[26] M. Maravanyika and N. Dlodlo, "Selecting attributes for inclusion in an educational recommender system using the multi-attribute utility theory," in Communications in Computer and Information Science, 2018, vol. 942, pp. 239–252, doi: 10.1007/978-3-030-01535-0_18.

[27] M. Kaity and V. Balakrishnan, "Sentiment lexicons and non-English languages: a survey," Knowledge and Information Systems. 2020, doi: 10.1007/s10115-020-01497-6.

[28] N. Al-Twairesh, H. Al-Khalifa, and A. Al-Salman, "AraSenTi: large-scale twitter-specific Arabic sentiment lexicons," in Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, 2016, pp. 697–705.

[29] A. Pasha et al., "MADAMIRA: a fast, comprehensive tool for morphological analysis and disambiguation of Arabic," in Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14), 2014, pp. 1094–1101, [Online]. Available: http://www.lrec-conf.org/proceedings/lrec2014/pdf/593_Paper.pdf.

[30] M. Hu and B. Liu, "Mining and summarizing customer reviews," in Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2004, pp. 168–177, doi: 10.1145/1014052.1014073.

[31] T. Wilson, J. Wiebe, and P. Hoffmann, "Recognizing contextual polarity in phrase-level sentiment analysis," in Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing, 2005, pp. 347–354, [Online]. Available: https://www.aclweb.org/anthology/H05-1044.pdf.

[32] F. H. H. Mahyoub, M. A. Siddiqui, and M. Y. Dahab, "Building an Arabic sentiment lexicon using semi-supervised learning," Journal of King Saud University - Computer and Information Sciences. vol. 26, no. 4, pp. 417–424, 2014, doi: 10.1016/j.jksuci.2014.06.003.

[33] M. Maamouri, D. Graff, B. Bouziri, S. Krouna, A. Bies, and S. Kulick, "LDC standard Arabic morphological analyzer (SAMA) version 3.1," Philadelphia, 2010. [Online]. Available: https://catalog.ldc.upenn.edu/.

[34] A. Esuli and F. Sebastiani, "SentiWordNet: a high-coverage lexical resource for opinion mining *," Italy, 2007. [Online]. Available: http://nmis.isti.cnr.it/sebastiani/Publications/2007TR02.pdf.

[35] S. Baccianella, A. Esuli, and F. Sebastiani, "SENTIWORDNET 3.0: an enhanced lexical resource for sentiment analysis and opinion mining," 2010, [Online]. Available: http://www.lrec-conf.org/proceedings/lrec2010/pdf/769_Paper.pdf.

[36] C. Fellbaum et al., "Introducing the Arabic wordnet project," in Proceedings of the 3rd Global Wordnet Conference, P. Sojka, K.-S. Choi, C. Fellbaum, and P. T. J. M. Vossen, Eds. Jeju Island, Korea, 2006.

[37] R. M. Duwairi, R. Marji, N. Sha'Ban, and S. Rushaidat, "Sentiment analysis in Arabic tweets," 2014, doi: 10.1109/IACS.2014.6841964.

[38] M. Abdul-Mageed, M. Diab, and S. Kübler, "SAMAR: subjectivity and sentiment analysis for Arabic social media," Computer Speech and Language. vol. 28, no. 1, pp. 20–37, 2014, doi: 10.1016/j.csl.2013.03.001.

[39] B. Liu, M. Hu, and J. Cheng, "Opinion observer: analyzing and comparing opinions on the web," 2005, doi: 10.1145/1060745.1060797.

[40] D. Kang and Y. Park, "Review-based measurement of customer satisfaction in mobile service: sentiment analysis and VIKOR approach," Expert Systems with Applications. vol. 41, no. 4 PART 1, pp. 1041–1050, 2014, doi: 10.1016/j.eswa.2013.07.101.

[41] Y. M. Li and T. Y. Li, "Deriving market intelligence from microblogs," Decision Support Systems. vol. 55, no. 1, pp. 206–217, Apr. 2013, doi: 10.1016/j.dss.2013.01.023.

[42] I. Chaturvedi, E. Cambria, R. E. Welsch, and F. Herrera, "Distinguishing between facts and opinions for sentiment analysis: survey and challenges," Information Fusion. vol. 44, pp. 65–77, Nov. 2018, doi: 10.1016/j.inffus.2017.12.006.

[43] A. Kaur and V. Gupta, "A survey on sentiment analysis and opinion mining techniques," Journal of Emerging Technologies in Web Intelligence. vol. 5, no. 4. pp. 367–371, Nov. 2013, doi: 10.4304/jetwi.5.4.367-371.

[44] M. Korayem, D. Crandall, and M. Abdul-Mageed, "Subjectivity and sentiment analysis of Arabic: A Survey," in Communications in Computer and Information Science, 2012, vol. 322, pp. 128–139, doi: 10.1007/978-3-642-35326-0_14.

[45] H. P. Patil and M. Atique, "Sentiment analysis for social media: a survey," in 2015 2nd International Conference on Information Science and Security (ICISS), 2015, pp. 1–4, doi: 10.1109/ICISSEC.2015.7371033.

[46] V. S. Jagtap and K. Pawar, "Analysis of different approaches to sentence-level sentiment classification," International Journal of Scientific Engineering and Technology. vol. 2, no. 3, pp. 164–170, 2013.

[47] A. D'Andrea, F. Ferri, and P. Grifoni, "Approaches, tools and applications for sentiment analysis implementation," International Journal of Computer Applications. vol. 125, no. 3, pp. 975–8887, 2015.

[48] A. B. Pawar, M. A. Jawale, and D. N. Kyatanavar, "Fundamentals of sentiment analysis: Concepts and methodology," in Studies in Computational Intelligence, vol. 639, Springer Verlag, 2016, pp. 25–48.

[49] E. Aydoğan and M. A. Akcayol, "A comprehensive survey for sentiment analysis tasks using machine learning techniques," in 2016 International Symposium on INnovations in Intelligent SysTems and Applications (INISTA), 2016, pp. 9–19, doi: 10.1109/INISTA.2016.7571856.

[50] Y. Jararweh, M. Al-Ayyoub, and E. Benkhelifa, "Advanced Arabic natural language processing (ANLP) and its applications: introduction to the special issue," Information Processing and Management. vol. 56, no. 2. Elsevier Ltd, pp. 259–261, Mar. 01, 2019, doi: 10.1016/j.ipm.2018.09.003.

[51] S. R. El-Beltagy and A. Ali, "Open issues in the sentiment analysis of Arabic social media: A case study," in 2013 9th International Conference on Innovations in Information Technology, IIT 2013, 2013, pp. 215–220, doi: 10.1109/Innovations.2013.6544421.

[52] N. A. Abdulla, N. A. Ahmed, M. A. Shehab, M. Al-Ayyoub, M. N. Al-Kabi, and S. Al-rifai, "Towards improving the lexicon-based approach for arabic sentiment analysis," International Journal of Information Technology and Web Engineering. vol. 9, no. 3, pp. 55–71, Jul. 2014, doi: 10.4018/ijitwe.2014070104.

[53] U. R. Hodeghatta, "Sentiment analysis of Hollywood movies on twitter," in Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2013, 2013, pp. 1401–1404, doi: 10.1145/2492517.2500290.

[54] M. Jokhio, N. A. Mahoto, S. Jokhio, and M. S. Jokhio, "Detecting tweet-based sentiment polarity of plastic surgery treatment," Mehran University Research Journal of Engineering and Technology. vol. 34, no. 4, pp. 403–412, 2015, [Online]. Available: https://www.researchgate.net/publication/283264994.

[55] U. Naseem and K. Musial, "DICE: deep intelligent contextual embedding for twitter sentiment analysis," in Proceedings of the International Conference on Document Analysis and Recognition, ICDAR, Sep. 2019, pp. 953–958, doi: 10.1109/ICDAR.2019.00157.

[56] U. Naseem, S. K. Khan, I. Razzak, and I. A. Hameed, "Hybrid words representation for airlines sentiment analysis," in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2019, vol. 11919 LNAI, pp. 381–392, doi: 10.1007/978-3-030-35288-2_31.

[57] U. Naseem, I. Razzak, K. Musial, and M. Imran, "Transformer based deep intelligent contextual embedding for twitter sentiment analysis," Future Generation Computer Systems. vol. 113, pp. 58–69, Dec. 2020, doi: 10.1016/j.future.2020.06.050.

[58] G. Yoo and J. Nam, "A hybrid approach to sentiment analysis enhanced by sentiment lexicons and polarity shifting devices," in The 13th Workshop on Asian Language Resources, 2018, pp. 21–28, [Online]. Available: https://hal.archives-ouvertes.fr/hal-01795217.

[59] H. K. Aldayel and A. M. Azmi, "Arabic tweets sentiment analysis - a hybrid scheme," Journal of Information Science. vol. 42, no. 6, pp. 782–797, 2015, doi: 10.1177/0165551515610513.

[60] J. Jia, G. W. Fischer, and J. S. Dyer, "Attribute weighting methods and decision quality in the presence of response error: a simulation study the full cost of electricity view project," Journal of Behavioral Decision Making. 1998, doi: 10.1002/(SICI)1099-0771(199806)11:23.0.CO;2-K.

[61] J. R. Bettman, M. F. Luce, and J. W. Payne, "Constructive consumer choice processes," Journal of Consumer Research. vol. 25, no. 3, pp. 187–217, 1998, doi: 10.1086/209535.

[62] F. Pedregosa et al., "Scikit-learn: machine learning in Python," Journal of Machine Learning Research. vol. 12, pp. 2825–2830, 2011.