Evaluation and Enrichment of Arabic Sentiment Analysis

Sanjeera Siddiqui, Azza Abdel Monem and Khaled Shaalan

Abstract This paper presents an innovative approach that explores the role of lexicalization for Arabic sentiment analysis. Sentiment Analysis in Arabic is hindered due to lack of resources, language in use with sentiment lexicons, pre-processing of dataset as a must and major concern is repeatedly following same approaches. One of the key solution found to resolve these problems include applying the extension of lexicon to include more words not restricted to Modern Standard Arabic. Secondly, avoiding pre-processing of dataset. Third, and the most important one, is investigating the development of an Arabic Sentiment Analysis system using a novel rule-based approach. This approach uses heuristics rules in a manner that accurately classifies tweets as positive or negative. The manner in which a series of abstraction occurs resulting in an end-to-end mechanism with rule-based chaining approach. For each lexicon, this chain specifically follows a chaining of rules, with appropriate positioning and prioritization of rules. Expensive rules in terms of time and effort thus resulted in outstanding results. The results with end-to-end rule chaining approach achieved 93.9% accuracy when tested on baseline dataset and 85.6% accuracy on OCA, the second dataset. A further comparison with the baseline showed huge increase in accuracy by 23.85%.

Keywords Sentiment analysis · Opinion mining · Rule-based approach
Arabic natural language processing

S. Siddiqui (✉) · K. Shaalan
British University in Dubai, Block 11, 1st and 2nd Floor, Dubai International Academic City,
Dubai, UAE
e-mail: faizan.sanjeera@gmail.com

K. Shaalan
e-mail: khaled.shaalan@buid.ac.ae

A.A. Monem
Faculty of Computer and Information Sciences, Ain Shams University,
Abbassia, Cairo 11566, Egypt
e-mail: azza_monem@hotmail.com

K. Shaalan
School of Informatics, University of Edinburgh, Edinburgh, UK

1   Introduction

The mission of Sentiment Analysis is to perceive the content with suppositions and mastermind them in a way complying with the extremity, which incorporates: negative, positive or nonpartisan. Organization's are taken to huge prestige by living up to the opinions from different people [1, 2]. Subjectivity and Sentiment Analysis characterization are prepared in four measurements: (1) subjectivity arrangement, to estimate on Subjective or Objective, (2) Sentiment Analysis, to anticipate on the extremity that could be negative, positive, or impartial, (3) the level in view of record, sentence, word or expression order, and (4) the methodology that is tailed; it could be standard based, machine learning, or half breed [3]. As stated by Liu [4], Sentiment Analysis is "Sentiment analysis, also called opinion mining, is the field of study that analyzes people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions towards entities such as products, services, organizations, individuals, issues, events, topics, and their attributes."

Arabic is a Semitic dialect, which is distinctive as far as its history, structure, diglossic nature and unpredictability Farghaly and Shaalan [5]. Arabic is broadly talked by more than 300 million individuals. Arabic Natural Language Processing (NLP) is testing and Arabic Sentiment Analysis is not a special case. Arabic is exceptionally inflectional [6, 7] because of the fastens which incorporates relational words and pronouns. Arabic morphology is intricate because of things and verbs bringing about 10,000 root [8]. Arabic morphology has 120 examples. Beesley [8] highlighted the hugeness of 5000 roots for Arabic morphology. No capitalization makes Arabic named substance acknowledgment a troublesome mission [9]. Free order of Arabic Language brings in additional challenges with regards to Sentiment Analysis, as the words in the sentence can be swapped without changing the structure and the meaning. Arabic Sentiment Analysis has been a gigantic center for scientists [10].

The target of this study is to examine procedures that decide negative and positive extremity of the information content. One of the critical result would be to recognize the proposed end-to-end principle binding way to deal with other dictionary based and machine learning-construct approaches in light of the chosen dataset.

The rest of this paper is organized as follows. Related work is covered in Sect. 2, Challenges to perform Arabic Sentiment Analysis is traced in Sect. 3, Data collection is covered in Sect. 4 followed by system implementation in Sect. 5. Section 6 covers evaluation and results, Sect. 7 depicts conclusion.

2 Related Work

To take a shot at Sentiment Analysis, the key parameter is the dataset. Late endeavors by Farra et al. [11] outlined the significance of crowdsourcing as an extremely fruitful technique for commenting on dataset. Sentiment corpora for Arabic brought into existence to prove the methods and perform experiments [12–16].

One of the extreme need to perform sentiment analysis is the availability of corpus. Few attempts are significant in literature illustrating the researchers zest to fill in the barren land of Arabic datasets to perform Arabic Sentiment Analysis. A freely available corpus brought into existence with the aid of manual approach by Al-Kabi et al. [17] covered five areas that is—Religion, Economy, Sport, Technology and Food-Life style, with each of these areas covering 50 subjects.

Al-Kabi et al. [17] collected reviews from Yahoo Maktoob with constraints including a clear discretion in the way the Arabic reviews written, with a mixture of Eygptian, English, MSA and Levantine dialect. Unlike Al-Kabi et al. [17] who have manually compiled the dataset, Farra et al. [18] used crowdsourcing to annotate dataset. Apart from freely available corpus, there has been significant work on different corpora but not shared openly to our knowledge. However some other corpus worked on includes Financial—[19, 20]; News—[21–23].

Other than the corpus, the other main player in the whole process of Arabic Sentiment Analysis is the Lexicon. Lexicons are the most desirable word list playing a key role in defining the text polarity. The lexicons can be available with the inclusion of adjectives. Adjectives are the best way to highlight polarity [24]. Few notable efforts found in literature where lexicons brought to life included, business based lexicons by [25], SentiStrength [26].

Some of the exemplar additions to the Arabic Sentiment Analysis include exploration on different methods in different domains however, piled up with limitations. Likewise, the work by Mountassir et al. [27] on documents with 2,925 reviews in Modern Standard Arabic wherein no experiments performed to judge their proposed method. With the news webpages and social reviews, Al-Kabi et al. [28] worked at sentence level by dealing with colloquial Arabic with a limitation of tiny dataset. However, Elarnaoty et al. [21] worked on document level with only center on news publications.

Bolster Vector Machine classifier accomplished 72.6% precision on twitter dataset of 1000 tweets [29]. The record level assessment investigation utilizing a joined methodology comprising of a dictionary and Machine Learning approach with K-Nearest Neighbors and Maximum Entropy on a blended area corpus involving training, legislative issues and games achieved an F-measure of 80.29% [30]. Shoukry and Refea [29] achieved a precision of 72.6% with the corpus based method. The vocabulary and estimation examination device with an exactness of 70.05% on tweeter dataset also, 63.75% on Yahoo Maktoob dataset [12].

With a mixed approach that is Lexical and Support Vector Machine classifier created 84.01% exactness [31]. 79.90% exactness was shown with Hybrid methodology which involved lexical, entropy and K-closest neighbor [32]. Shoukry and Rafea [33] independently sent two methodologies, one being Support Vector Machine accomplished a precision of 78.80% and other one being Lexical with an exactness of 75.50%.

3 Challenges Arabic Sentiment Analysis

In the first place the challenges with regards to sentiment analysis in Arabic includes a principal that has a quick effect is subjectivity portrayal. In light of the very truth of a trim distinction in a sentence being subjective or objective gets new challenges. Various diverse challenges are found in Sentiment Analysis. If a negative presumption contains a segment from the positive vocabulary, it is general explored as positive limit wherein it is negative. A word found as a section of both positive and negative vocabularies obtains new challenges in the furthest point undertaking. For example, "قرف با" (it makes me sick or disgusting) was found in both negative and positive reviews; finally, joining positive and negative vocabularies. A couple people create rude comments or unmistakable style of commenting using negated sentences in a productive review thusly gets new troubles. For example, "احسن من هيك رئيس وزراء ما في" (There is no Prime Minister better than this one).

As spoke to by Pang and Lee [34], positive words in one zone may hold a substitute furthest point in another space. For example, a word like "بارد" (cold) were found in positive and negative reviews holding specific limit. Distinctive challenges joins, beyond what many would consider possible to 140 characters people winds up using short structures, semantic mistakes in the tweets and specifically usage of slang tongue [31]. Following are few challenges hindering the natural language processing in Arabic.

3.1 Encoding

Windows CP-1256 and Unicode availability, one can read, compose and extract substance written in Arabic. Issues do rise in the midst of the preparation work by some different tasks, as needs be the use of transliteration where the Roman letters and Arabic letters mapping are possible. Subsequently, keeping in mind the end

goal to decide this issue, various researchers use transliteration as the pre-processing ready stride [35].

3.2 Sentiment Analysis Impacted Due to Unavailability of Punctuations

A great degree smart parts/features shared by Arabic, it gets new troubles. One of the key challenge found in Arabic NLP is the nonattendance of strict and rigid standards with reference to complement (punctuation), get challenges for stick point sentence confines in Modern standard Arabic [36]. Same basic is found in Dialect Arabic, with no fundamentals to manage boundaries of a sentence. No complements (punctuations) are trailed by a huge bit of the all-inclusive community except for development of a full stop toward the end, which even is habitually overlooked. With no highlights, enormous effect is found in most of the NLP assignments including Arabic Sentiment Analysis settling on the decision of right sentences troublesome from a given substance.

3.3 Excess Resources Required

Turney [37] highlighted the deficiency of tools and resources in general for Arabic resulting in excess challenges for community doing research in Arabic NLP. The tools and resources forms the assets for Arabic NLP and are found to be the most lacking area when vocabularies and corpus are talked about in sentiment analysis. The vocabularies and corpus forms noteworthy part in Sentiment Analysis.

Because of the absence of culmination in the dictionary for positive and negative thwarts the investigation of Sentiment. Abdalah et al. [12] and Farra et al. [22], expressed regardless of analysts are investing vitality and time creating vocabularies and gathering tweets and audits framing dataset, yet these are not made accessible open for others to utilize and investigate, frustrated the development in Arabic Natural Language Processing with Sentiment Analysis in setting.

Table 1 Lexicons constructed are not complete

Paper	Method used	Lexicon build for domain	Own/translated
[35]	Unsupervised learning	Business reviews	Own
[36]	By translating English lexicons into Arabic	English lexicons for various reviews.	Translated
[35]	By labelling 4000 adjectives as negative, positive and neutral	4000 adjectives words irrespective of any particular domain	Own

Table 2 Corpus for different domains (not openly shared)

Domain	Paper
Financial	[24, 26]
Twitter tweets	[31]
Multi domain	[32, 36]
News corpus	[21, 22]
Web reviews	[29]

Table 1 delineates the vocabulary worked for a portion of the areas which is exceptionally constrained not covering numerous spaces and are not unreservedly accessible. Table 2 portrays the datasets worked for multi space yet are not transparently accessible also are extremely restricted.

3.4   Sarcastic Tamper

Wry method for speaking to Arabic suppositions or opinions brings about huge difficulties for the framework to segregate a survey as positive or negative. The snide obstruction causes a gigantic harm to the judgment of general extremity. Due to sarcasm a negative audit could be anticipated as positive and the other way around.

An example sarcastic tweet:

مجرد مشاهد لصعود وهبوط الفك أوه نعم، وأنا أتفق أنه في بعض الأحيان أنا لا أستمع لما تقوله، وأنا ''

- Oh yes, I agree that sometimes I don't listen what you say, I just watch your jaw go up and down". Table 3, depicts some of the sarcastic comments which creates the polarity interpretation more cunning task.

Table 3 Sarcastic tamper

Sarcastic comments	
Arabic	English translation
أحيانا أنا لا أعرف إذا كانت اصرخ عليك أو أشفق عليك	Sometimes I don't know whether to shout at you or pity you
أنت مضحك بحيث انك تجعل الجميع يضحكون عليك وليس معك	You are so funny that you are making everyone laugh at you, not with you
الفستان الذي كنت ترتديه يبدو لطيفا ولكن ليس عليك	The dress that you are wearing looks nice, but not on you
أحيانان أحتاج لشيء لا يمكن أن يعطيه أحد إلا أنت وهو أن تغلق فمك	Sometimes I need something that no one can give except you, which is your mouth shut

3.5 One Word Represents Two Polarities

In Arabic, addition or deletion of one word turns the sentence into opposite polarity. In the example A: "أنا أحب المدرسة - I like school" and "أنا لا أحب المدرسة - I do not like school". In like manner, a bit much a similar word constantly tends to give a negative significance.

In example B: "لا أحد يصعد الشجرة - No one climbed the tree" contains the word "لا - No" shows up in this sentence however does not pass on any negative assessment. Thus this arrangement of disagreements with single word speaking to two unique assumptions in two varied sentences brings about real concerns.

While fulfilling a negative supposition, one can fulfill "لا - No" as negative estimation by putting the word in negative vocabulary, however neglects to fulfill case B by breaking down the announcement as negative however it is definitely not. Table 4 portrays cases of the words changing sentence extremity. Table 5 delineates the representation for cases on the difficulties that one negative word can convey to the Sentiment Analysis errand in Arabic Natural dialect preparing.

Table 4 One word changing polarity—examples with challenges faced

Positive sentiment word	لذيذ — indicating positive polarity.	سعيد — happy
Positive review	الطعام لذيذ - The food is delicious	أنت سعيد في عملك — You are happy with your work
Addition of one word	ليس Converts the polarity to negative	Addition of لست converts the polarity to negative
Negative review	الطعام ليس لذيذ - The food is not delicious	وأنت لست سعيدا في عملك — You are not happy with your work

Table 5 Challenges to be dealt with

Challenges to tackled both negative and positive polarities	Brings in challenges due to the addition of لذيذ – delicious in positive lexicon, hence the negative sentiment containing the word لذيذ – delicious makes the system identifies the negative sentiment as positive. If "ليس – not" not tackled properly so as to satisfy negative polarity	Brings in challenges due to the addition of "سعيد – happy" in positive lexicon, hence the negative sentiment containing the word "سعيد – happy" makes the system identify the negative sentiment as positive, if "ليس – not" not tackled properly so as to satisfy negative polarity

3.6 Indifferent Writing Style

Arabic audits for the most part found on various locales have a reasonable assorted qualities in composing style with Egyptian and Gulf Arabic regularly utilized. A large portion of the Arabic audits are not composed in standard arrangement and doesn't take after Modern Standard Arabic. Individuals compose distinctive vernacular in Arabic which are casual and in this manner includes the utilization of individual elocution and vocabularies [4]. Accordingly, the dictionary building procedure and control based drew nearer are basically tested with this casual writing styles leaving specialists with exhaustive investigation of surveys composed online to perform Sentiment Analysis. Table 6 delineates different written work styles

Table 6 Different writing style resulting in inexactness

Review/comment	Polarity	Negative/positive words	Challenges
انها ليست امرأة سيئة ـ She is not a bad woman	Positive	Negative words— not and bad	This type of writing style containing negative words but indicating positive polarity makes the job for the system a difficult task to handle. Not and bad due to the very nature of these words indicating their significance in negative reviews, joins the negative lexicon. When the system tries to tackle this positive sentiment, but due to the addition of negative words the polarity system gives as an output is negative but the correct polarity is positive
عليك أن تكون كريما ـ you have to be generous	Negative	Positive words— generous	This type of writing style with addition of positive words resulting in negative polarity brings in challenges wherein the system matches this with both positive and negative lexicon and indicates the polarity as positive though this sentence is negative, if need to be is not tacked to resolve this challenge
أنت لست انسانن جيدا ـ You are not a good human being	Negative	Positive word— good negative— not	System struggles to tackle the existence of words like good and not belonging to both positive and negative lexicon. Based on the way the system is designed, the overall polarity goes in favour of either positive or negative. If system indicate this as positive polarity then the answer is wrong, indicating that the classification lacks in handling mixed words in a sentiment
ليس هناك احد حزين في هذا البلد ـ No one is sad in this country	Positive	Negative words— no sad	This example though include two negative words but fails to satisfy the overall polarity if the system analyse this based on the negative lexicon list and identify this as negative review though the review is positive resulting in inexactness

alongside the difficulties they acquire to the framework to fulfill the negative and positive polarities.

3.7 Free Writing Style

Another key issue found to affect Arabic Natural Language Processing particularly in Sentiment Analysis, wherein individuals don't take after linguistic use, accentuations, and spellings. Individuals are all the time appear to overhaul their status and remark online without breaks, this may bring about an inclination to miss spell or write in rush missing the fundamental structure of composing a legitimate sentence. This makes the Sentiment Analysis prepare more mind boggling. Spelling slip-ups are high in online surveys and tweets [9].

Table 7 delineates couple of cases of free written work style with the spelling botches as greatest test.

3.8 Word Short Forms

Because of the cutoff points on numerous web-based social networking locales case Twitter, with respect to the quantity of words that could be composed, gets new difficulties for performing Natural Language Processing. The way the Arabic tweets are abbreviated by putting the short structures for the words impedes the System to discover the match for the feeling in the Lexicons. Like "غرامة - Fine" is frequently trimmed and composed as "F9" [31].

Table 7 Challenges faced due to spelling mistakes in free writing style

	Lexicon	Review in Arabic	Review in english	Challenges due to the spelling mistakes
Actual sentence	أفضل	أنت أفضل شخص في المبيعات	You're the best sales person	Actual polarity is positive, due to the word misspelled as ضل for أفضل, turns the whole polarity of the review into negative
Spelling mistake	ضل spelled incorrectly for أفضل	أنت ضل شخص في المبيعات	You are lost sales person.	

3.9 Same Word Usage for Both Polarities

At times individuals answer to a remark or audit or tweet through single word or expressions which can bolster both positive and negative survey, subsequently falling in both positive and negative vocabularies. Not just the initial step that is the vocabulary building process gets to be distinctly testing, additionally the framework building turns out to be tremendously testing. Colossal examination is required in this circumstance to deal with such words. Words like "okay, fine, alright, all right, very well – حسنا ", "أنا أتفق – I agree", " اوافق لا انا - I do not agree" are frequently observed to bolster assumptions in negative and additionally positive. Building framework understanding on when to counter to these supporting or not supporting words or expressions as negative or positive is a noteworthy hiccup. Table 8 delineates a response to twitter tweets and the difficulties these reactions conveys to the framework to anticipate the extremity of the reaction.

Table 8 Person to person response challenge

	Twitter tweet	Polarity	Challenges
Person-1	ساره دائما تتظاهربأنها جيدة جدا في الطبخ لكنها ليست جيدة في الطبخ. – Sara always pretend to be very good in cooking but she is not a good cook	Negative	This is the challenging phase wherein making the system understand when to give a positive polarity if the word "حق – right" is found in a tweet, as this word belongs to the positive lexicon, but in the current case it supports the negative polarity. Hence illustrating a negative sentiment by supporting a negative statement
Person-2 responding to person-1	انت دائما على حق – You are always right	Supporting a negative opinion—negative polarity	
Person-3	اعتقد انها لا تعرف الف باء الطبخ - I think she does' not know the abc of cooking	Negative comment	Person-3 is negative sentiment. This negative sentiment is not supported by person-4 instead gives a positive response by not agreeing to person-3. Due to the addition of "أنا لا أتفق معك - I do not agree", brings in new challenges as the system will predict this as negative, due to "I do not agree - أنا لا أتفق معك" which falls under negative sentiment and the system will consider this to be a negative review but in actual this is a positive response to a negative review
Person-4 responds to person-3	أنا لا أتفق معك - I do not agree with you	Positive response do not supporting a negative sentiment	

4 Data Collection

The dataset utilized as a part of this paper is Twitter Tweets and film surveys. The dataset is taken from [12, 38]. These datasets are utilized as these are accessible, rich and enough to reached a conclusion, also electronic assets, for example, vocabulary is given.

Abdulah et al. [12] dataset contains 1000 positive tweets and 1000 negative tweets with length ranging from one word to sentences. 7189 words in positive tweets and 9769 words in negative tweets. The tweets were manually collected belong to Modern Standard Arabic and Jordanian Dialect, which covers Levantine language family. The months-long segregation procedure of the tweets was physically led by two human specialists (local speakers of Arabic).

OCA corpus, termed as Opinion Corpus for Arabic, was presented by [38]. This corpus contains total 500 opinions, of which there are 250 positive opinions and 250 negative opinions. The procedure followed by Rushdi et al. [38] included gathering surveys from a few Arabic online journal locales and site pages utilizing a straight-forward bash script for slithering. At that point, they expelled HTML labels and unique characters, and spelling mix-ups were adjusted physically. Next, a preparing of each survey was done, which included tokenizing, evacuating Arabic stop words, and stemming and sifting those tokens whose length was under two characters. In their trials, they have utilized the Arabic stemmer of RapidMiner and the Arabic stop word list. At last, three distinctive N-gram plans were created (unigrams, bigrams, and trigrams) and cross validation was used to assess the corpus for which they have achieved 90.6% accuracy.

The vocabulary utilized as a part of this paper contains the dictionary used by [12], which included opinions, named substances and a few haphazardly set words. In view of the circumspection of reiteration of the words found in negative or positive reviews and their arrangement, they were incorporated into both the run-down that is positive and negative lists.

5 Implementation of Arabic Sentiment Analysis

This paper is a significant extension of [11]. Siddiqui et al. [11] research introduced a system, which contains only two type of rules—"equal to" and "within the text rules", so as to examine whether the tweet is either negative or positive. The rules include a 360° coverage with an improvised segment that is end to end rule chaining principle: (1) in the middle, we termed as "within the text", (2) at the boundary we termed as either "ending with the text" or "beginning with the text", and (3) full coverage, we termed as "equal to the text". Figure 1 delineates the 360° rules coverage. The End-to-End mechanism with rule chaining approach introduced in this paper, includes the chaining of rules based on the positioning of the polarities in the tweets. The key underlying base ground factors which helped us formulate
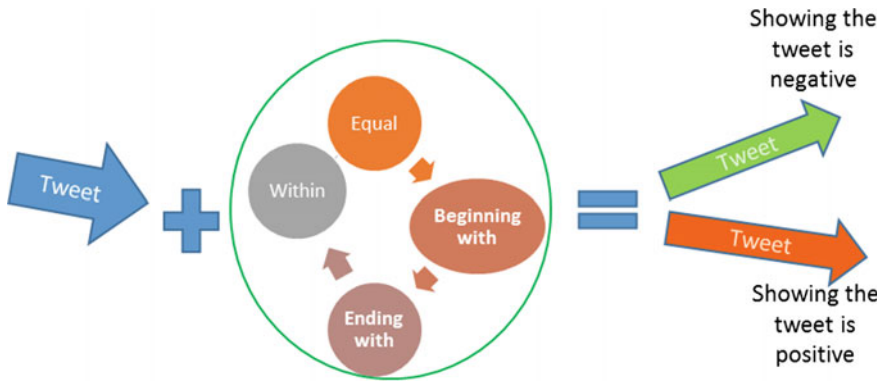
Fig. 1 A 360° coverage of rules to the input tweet

appropriate rules includes analysis of the tweets and the extension of positive and negative lexicons. The analysis of tweets resulted in identifying relations pertaining to words which were either disjoint, intersected or coexisted.

The words which were disjoint that is completely indicating either positive or negative polarity were included in their respective lexicons. The words which intersected that is the ones which were found to be common in both negative and positive reviews were included in positive as well as negative lexicons. The words which coexisted at the same place in the negative and positive reviews, that is the ones which appeared at the beginning or ending were placed in either positive or negative lexicon, based on the highest frequency of the word in the respective reviews.

Rules handling intersection with the end-to-end chaining mechanism: As an example consider the following positive tweet "اوقف القرار للحفاظ على الوطن" (the decision was suspended for protecting the motherland), the word "اوقف" (suspended) appeared in the beginning of this tweet and the same word was found in the negative tweets. Hence, this set of situations was handle with the very use of positioning and chaining of rules. The steps involved to achieve the positioning and chaining of rules includes:

Rules Formation: With the logical discretion of the word "اوقف" (suspended) being seen at the "beginning of" positive tweets and "within the text" for the negative tweets, the rules thus formed were "beginning with" for positive tweets and "within the text" for negative tweets.

With the correct positioning and chaining of rules this problem was resolved. In the current example "beginning with" rule needs to be positioned and chained in an orderly fashioned with the rule "within the text" so as to satisfy both positive and negative reviews. So, the rule "beginning with" was chained with the rule "within the text" for the word "اوقف" (suspended) by positioning the rule "beginning with" first followed by "within the text" rule. Hence, the rules are chained and positioned

for the words which are found to repeat themselves in both positive and negative tweets.

Rules handling coexistence: The lexicons which are not seen to repeat themselves in either positive or negative tweets where handled with the rule—"within the text". For example, "الحرامية" (Thieves) is set for search "within the text" "الحرامية" (Thieves), is majorly found within the text in negative review rather than in positive ones. Hence based on the frequency of "الحرامية" (Thieves) the rule is set. Example tweet: " مسؤول عن الحرامية يتفنجرو بمصاري البلد و انا اشحد انا مش " (I am not responsible for that thieves spend lavishly from the funds of the country and I beg).

Rules handling disjoint: The cases wherein the words were not repeated in positive and negative cases and were found to have their significance at the end of the tweet, were handled using "ending with" rule. In the following positive tweet example " من يكذب يكتب عند الله كذابا – اما من يتحرى الصدق يكتب عند الله صادقا تذكر "—(Remember who continues to speak falsehood he is recorded with Allah as a great liar—but who persists in speaking the truth he is recorded with Allah as an honest man).

The word "صادقا" (Honest) appeared at the end of this tweet. Likewise, "صادقا" (Honest) was found to appear at the end in majority of positive reviews. Hence, the "end with" rule is set to search for "صادقا" indicating the system that if the review ends with the word "صادقا" (Honest) then it should be considered as a positive tweet.

Derivation of rules: Table 9 depicts the skeletal examples of derived rules. Conditional search includes two key phases, one includes a condition which checks on the entered tweet mapping with the rules and the second phase is the color coded output which changes its font color to—"Green fill with Green fill text" for negative tweets and "Light red fill with light red fill text" for positive tweets.

Table 9 Derivation of lexicalized rule

Rule type		Derivation of rules
(1) Search within the text	A	If the entered tweet/review contains the word "__" then fill the text with—"Green fill text" for negative tweets
	B	If the entered tweet/review contains the word "__" then fill the text with—"Light fill text" for positive tweets
(2) Ending with the text	A	If the entered tweet/review ends with the word "__" then fill the text with—"Green fill text" for negative tweets
	B	If the entered tweet ends with the word "__" then fill the text with —"Light red fill text" for positive tweets
(3) Beginning with the text	A	If the entered tweet begins with the word "__" then fill the text with—"Green fill text" for negative tweets
	B	If the entered tweet begins with the word "__" then fill the text with—"Light red fill text" for positive tweets
(4) Equal to the text	A	If the Entered Tweet equals to the word "__" Then Fill the text with—"Green fill text" for positive tweets
	B	If the Entered Tweet equals to the word "__" then Fill the text with —"Light red fill text" for negative tweets

With reference to rule 1.A in Table 9 in the primary column, if the word showed up inside of the content in a positive tweet. With reference to manage 2.B or 3.B in this table, if the same word showed up toward the end or toward the starting in the negative tweet, then situating and anchoring of these two rules are finished. For this case the positive principle was situated and fastened underneath the negative guideline. The principle "starting with" or "finishing with" not at all like "inside of the content" search for that word in the first place or end which will tag the tweet as negative. In the event that the situation of guideline is turned around then "inside of the content" as situated and affixed before the tenets "starting" or "consummation" with, when a negative tweet is gone through this, the tweet will be labeled positive. As this is a chained approach for a specific word, this checks its importance at the ending or starting rule, on the off chance that it doesn't have a place with that then it goes through within the text.

End-to-end mechanism with rule chaining approach: After precisely executing the rules taking into account the fulfillment of disjoint word(s) which existed together or basic words in negative and positive surveys, the guideline anchoring was dealt with. End-to-end component with rule chaining approach fulfills a word which has a place with a positive and negative tweet regardless of its real extremity. Consequently, a negative word in a positive audit and a positive word in a negative tweet is fulfilled through apt chaining as examined beneath with the guide of case from negative and positive tweets. Example 1 Rules in use—Rule A and Rule B.

Rule A in Table 9: In the negative tweet "ليست هناك حاجة لحماية وطننا من أيدي المنافقين الأردنيين", there is no need to protect our motherland from hands of hypocrites Jordanian, contains the word "المنافقين" (hypocrites) within the text. Rule B in Table 9: In the positive tweet "اللهم لا تجعلنا من المنافقين" (O Allah! Place us not with the people who are hypocrites) contains the word "المنافقين" (hypocrites) at the last. If Rule A is not chained with Rule B, then only one of the rule will be satisfied. To satisfy both the rules, that is to correctly identify negative and positive polarity, Rule A is chained with Rule B by positioning Rule A Below Rule B, so as to allow the search to first visit the end rule first, then the within the text rule.

Example 2 with explanation on how the system works: For example, the word "تشائم" (pessimism) was found to be part of both positive and negative reviews with a slight variation. In the positive reviews appears only in the middle whereas in the negative review "تشائم" (pessimism) was found to appear at the beginning. Hence the rules thus were created covering "within the text" (refer rule 1.A in Table 9) and "beginning with" (refer rule 2.B in Table 9), but the positioning was varied. By positioning the rule "beginning with" first and then the rule "within the text" helped in satisfying both positive and negative reviews.

As the system checked for the word beginning with "تشائم" (pessimism) in the entered review and if found then the review was tagged as negative. Likewise, the system proceeded with the entered review containing the word "تشائم" (pessimism), if the word "تشائم" (pessimism) was not found at the beginning then the search proceeded further and identified it as positive.

6 Evaluation and Results

To quantify the change also the nature of being trusted and had faith in, assessment plays an essential part. Cross-Validation and accuracy information are regularly used to assess the outcomes in estimation investigation. The exactness measures—Precision, Recall and Accuracy, which are generally being used was conveyed to measure the execution of the instruments utilized as a part of both the analyses. Precision, Recall and Accuracy were utilized to look at the outcomes by [12, 38, 33]. The condition is as per the following:

Precision = TP/(TP + FP)

Recall = TP/(TP + FN)

Accuracy = (TP + TN)/(TP + TN + FP + FN)

where:

TP—True Positive, all the tweets which were characterized accurately as positive

TN—True Negative, all the tweets which were accurately named negative

FP—False Positive, all the tweets which were mistakenly named positive

FN—False Negative, all the tweets which were mistakenly delegated negative

Results: The results fuse the relationship of the impressive number of tests coordinated in this paper. To do the connection, the accuracy of the extensive number of tests are used. Siddiqui et al. [11] system and System 2 (introduced in this paper) were attempted on [12, 38] dataset. Table 10, obviously takes after the outperformance of rules made in Siddiqui et al. [11] with enormous accuracy for Abdulah et al. [12] dataset when contrasted with the results on [38] OCA dataset.

System 1 Versus System 2 Results Comparison: Clear importance in expansion in accuracy in System 2 is seen for Abdulah et al. [12] and OCA dataset. The examination of Siddiqui et al. [11] System 1 and our System 2 doubtlessly answers that the end to end rule chaining improved the performance of sentiment analysis.

Siddiqui et al. [11]-system 1 bound to limits with two norms sort with no attaching and System 2 variation of System 1 with authoritative and reasonable arranging of the standards. System 2 ended up being remarkable for both the datasets. 93.9% accuracy for [12] dataset wherein for OCA dataset 85.6% accuracy was measured. Still the accuracy for our System 2 when tested on Abdulah et al. [12] was high with 8.3% more exactness than OCA dataset. Recall is high for both

Table 10 Comparison of system 1 [11] versus system 2

	System 1		System 2	
	Abdulah et al. [12] dataset	Rushdi et al. [38] OCA dataset	Abdulah et al. [12] dataset	Rushdi et al. [38] OCA dataset
Precision	87.4	50.4	92.5	77.4
Recall	93.3	97.6	95.8	100
Accuracy	89.6	50.1	93.9	85.6

the datasets with 3.3% for Abdulah et al. [12] and 22.6% for Rushdi et al. [38] as very less number of tweets are mistakenly delegated negative.

7 Conclusion

This paper beats the vocabulary building process through the fitting position of words too not barring the basic words found in both the tweets for the vocabularies. The outperformance of rule chaining approach that is System 2 brought about 23.85% in results when contrasted with Abdulah et al. [12] vocabulary based methodology. The incorporation of normal words in light of the examination of tweets in the negative and positive vocabulary list upgraded the general result when contrasted with the gauge dataset. Last yet not the minimum, the situating of principles has a gigantic effect to the rule based methodology as proper situating brought about fulfilling words which were observed to be regular in both negative and positive dictionaries. By and by, the end-to-end rule chaining methodology was the most difficult and costly regarding time and exertion, yet adds to the headways in the cutting edge for Arabic Sentiment Analysis, through the organized set standards and through the right utilization of various principles including "contains content", "equivalent to", "starting with" and "finishing with".

In reality, displayed the recently created assessment investigation framework—System 2, which beat in both arrangements of examinations when contrasted with [12]. System 2 with guidelines reached out to cover all territories was demonstrated to expand the exactness of OCA corpus by 39.8 and 4.3% accuracy for Abdulah et al. [12] when contrasted with System 1's principles of Siddiqui et al. [11]. Thus, starting a wakeup require every one of the scientists to redirect their enthusiasm to lead based methodology. The unmistakable hugeness in results along these lines acquired through the tenets made makes the principle based methodology the most alluring methodology.

References

1. Feldman, R.: Techniques and applications for sentiment analysis. Commun. ACM **56**(4), 82–89 (2013)
2. Taboada, M., Brooke, J., Tofiloski, M., Voll, K., Stede, M.: Lexicon-based methods for sentiment analysis. Comput. Linguist. **37**(2), 9–27 (2011)
3. Korayem, M., Crandall, D., Abdul-Mageed, M.: Subjectivity and sentiment analysis of Arabic: a survey. In: Hassanien, A.E., Salem, A.-B.M., Ramadan, R., Kim, T.-H. (eds.) AMLTA 2012. CCIS, vol. 322, pp. 128–139. Springer, Heidelberg (2012)
4. Albalooshi, N., Mohamed, N., Al-Jaroodi, J.: The challenges of Arabic language use on the internet. In: 2011 International Conference for Internet Technology and Secured Transactions (ICITST), pp. 378–382. IEEE (2011)

5. Farghaly, A., Shaalan, K.: Arabic natural language processing: challenges and solutions. ACM Trans. Asian Lang. Inf. Process. (TALIP) **8**(4), 1–22 (2009). The Association for Computing Machinery (ACM)

6. Hammo, B., Abu-Salem, H., Lytinen, S., Evens, M.: QARAB: a question answering system to support the Arabic language. In: The Proceedings of Workshop on Computational Approaches to Semitic Languages, ACL 2002, Philadelphia, PA, pp. 55–65, July 2002

7. Syiam, M., Fayed, Z., Habib, M.: An intelligent system for Arabic text categorization. Int. J. Intell. Comput. Inf. **6**(1) (2006)

8. Darwish, K.: Building a shallow Arabic morphological analyzer in one day. In: Proceedings of the ACL-2002 Workshop on Computational Approaches to Semitic languages, pp. 1–8. Association for Computational Linguistics, July 2002

9. Khaled, S.: A survey of Arabic named entity recognition and classification. Comput. Linguist. **40**(2), 469–510 (2014)

10. Muhammad, A.-M., Kübler, S., Diab, M.: Samar: a system for subjectivity and sentiment analysis of Arabic social media. In: Proceedings of the 3rd Workshop in Computational Approaches to Subjectivity and Sentiment Analysis, pp. 19–28. Association for Computational Linguistics (2012)

11. Siddiqui, S., Monem, A.A., Shaalan, K.: Sentiment analysis in Arabic. In: Métais, E., Meziane, F., Saraee, M., Sugumaran, V., Vadera, S. (eds.) Proceedings of the 21st International Conference on the Application of Natural Language to Information Systems (NLDB 2016). LNCS. Springer, Heidelberg (2016)

12. Abdulla, N.A., Ahmed, N.A., Shehab, M.A., Al-Ayyoub, M., Al-Kabi, M.N., Al-rifai, S.: Towards improving the lexicon-based approach for Arabic sentiment analysis. Int. J. Inf. Technol. Web Eng. (IJITWE) **9**(3), 55–71 (2014)

13. Abdulla, N., Al-Ayyoub, M., Al-Kabi, M.: An extended analytical study of Arabic sentiments. Int. J. Big Data Intell. **1**(2), 103–113 (2014)

14. Abdulla, N., Ahmed, N., Shehab, M., Al-Ayyoub M.: Arabic sentiment analysis: lexicon-based and corpus-based. In: Proceedings of IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies, Amman, Jordan, pp. 1–6 (2013)

15. Al Shboul, B., Al-Ayyoub, M., Jararweh, Y.: Multi-way sentiment classification of Arabic reviews. In: Proceedings of the 6th International Conference on Information and Communication Systems, Amman, Jordan (2015)

16. Al-Ayyoub, M., Bani-Essa, S., Alsmadi, I.: Lexicon-based sentiment analysis of Arabic tweets. Int. J. Soc. Netw. Min. **2**(2), 101–114 (2015)

17. Al-Kabi, M.N., Al-Ayyoub, M.A., Alsmadi, I.M., Wahsheh, H.A.: A prototype for a standard Arabic sentiment analysis corpus. Int. Arab J. Inf. Technol. (IAJIT), **13** (2016)

18. Farra, N., McKeown, K., Habash, N.: Annotating targets of opinions in Arabic using crowdsourcing. In: ANLP Workshop 2015, p. 89 (2015)

19. Ahmad, K.: Multi-lingual sentiment analysis of financial news streams. PoS, 001 (2006)

20. Almas, Y., Ahmad, K.: A note on extracting 'sentiments' in financial news in English, Arabic & Urdu. In: The Second Workshop on Computational Approaches to Arabic Script-based Languages, pp. 1–12 (2007)

21. Elarnaoty, M., AbdelRahman, S., Fahmy, A.: A machine learning approach for opinion holder extraction in Arabic language. arXiv:1206.1011 (2012)

22. Farra, N., Challita, E., Assi, R.A., Hajj, H.: Sentence-level and document-level sentiment mining for Arabic texts. In: Feldman, R. (ed.) 2010 IEEE International Conference on Data Mining Workshops (ICDMW), pp. 111 (2013). Techniques and applications for sentiment analysis. Commun. ACM **56**(4), 82–89 (2010)

23. Al-Kabi, M.N., Alsmadi, I.M., Gigieh, A.H., Wahsheh, H.A., Haidar, M.M.: Opinion mining and analysis for Arabic language. Int. J. Adv. Comput. Sci. Appl. (IJACSA) **5**(5), 181–195 (2014)

24. Benamara, F., Cesarano, C., Picariello, A., Recupero, D.R., Subrahmanian, V.S.: Sentiment analysis: adjectives and adverbs are better than adjectives alone. In: ICWSM (2007)

25. El-Halees, A.: Arabic opinion mining using combined classification approach (2011)

26. Elhawary, M., Elfeky, M.: Mining Arabic business reviews. In: 2010 IEEE International Conference on Data Mining Workshops (ICDMW), pp. 1108–1113. IEEE (2010)
27. Mountassir, A., Benbrahim, H., Berrada, I.: An empirical study to address the problem of unbalanced data sets in sentiment classification. In: 2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 3298–3303. IEEE (2012)
28. Al-Kabi, M., Gigieh, A., Alsmadi, I., Wahsheh, H., Haidar, M.: An opinion analysis tool for colloquial and standard Arabic. In: The Fourth International Conference on Information and Communication Systems (ICICS 2013), pp. 23–25 (2013)
29. Shoukry, A., Rafea, A.: Sentence-level Arabic sentiment analysis. In: 2012 International Conference on Collaboration Technologies and Systems (CTS), pp. 546–550. IEEE (2012)
30. Alaa, E.-H.: Arabic opinion mining using combined classification approach (2011)
31. Aldayel, H.K., Azmi, A.M.: Arabic tweets sentiment analysis–a hybrid scheme. J. Inf. Sci. (2015). doi:10.1177/0165551515610513
32. Beesley, K.R.: Arabic finite-state morphological analysis and generation. In: Proceedings of the 16th Conference on Computational Linguistics, vol. 1, pp. 89–94. Association for Computational Linguistics, August 1996
33. Shoukry, A., Rafea, A.: Preprocessing Egyptian dialect tweets for sentiment mining. In: The 4th Workshop on Computational Approaches to Arabic Script-Based Languages, pp. 47–56 (2012)
34. Pang, B., Lee, L.: Opinion mining and sentiment analysis. Found. Trends Inf. Retrieval **2**(1–2), 1–135 (2008)
35. ElSahar, H., El-Beltagy, S.R.: Building large Arabic multi-domain resources for sentiment analysis. In: International Conference on Intelligent Text Processing and Computational Linguistics, pp. 23–34. Springer (2015)
36. Diab, M., Hacioglu, K., Jurafsky, D.: Automatic processing of modern standard Arabic text. In: Arabic Computational Morphology, pp. 159–179. Springer, Netherlands (2007)
37. Turney, P.D.: Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. In: Proceedings of the Association for Computational Linguistics, pp. 417–424. Association for Computational Linguistics (2002)
38. Rushdi-Saleh, M., Martín-Valdivia, M.T., Ureña-López, L.A., Perea-Ortega, J.M.: OCA: opinion corpus for Arabic. J. Am. Soc. Inf. Sci. Technol. **62**(10), 2045–2054 (2011)