Automated Analysis of Comments on Press Articles in Websites (Journalist of Articles on Al-Ahram newspaper as study case)

Afra Arif [a]   ,   Abdul Monem S. Rahma [b]

[a] Department of Computer Science, University of Technology, Baghdad, Iraq. Email: afra.russel@yahoo.com

[b] Department of Computer Science, University of Technology, Baghdad, Iraq. Email: 110003@uotechnology.edu.iq

A R T I C L E   I N F O

A B S T R A C T

Due to the increasing writers of press articles on electronic publishing sites for printed and certified newspapers, the reader faced the problem of reaching his goal by accessing these sites, which led to the neglect of a large section of important publications of press articles for the writers. The provision of an automated measure to verify the positive and negative press articles of the writers based on the analysis of readers' comments on the press articles is a necessity to see the important writers of the press articles that is compatible with corpus generated by us for inference. the proposed system was applied on ten press articles by the writer "". thses ten articles also evaluated by the experts. The project achieved the previous target and achieved a success rate. The proposed system is evaluated based on four metrics (Precision, Recall, F-measure and Accuracy) . The effectiveness results obtain by this system was (Precision= 95%, Recall= 95%, F-measure=95% and Accuracy=94%).

MSC.

Corresponding author Afra Arif
Email addresses: afra.russel@yahoo.com
Communicated by Qusuay Hatim Egaar

_____

1. Introduction

Recently, there has been an increase in the number of press articles published in the newspaper sites on the internet, due to the emergence of many writers of press articles and a large number of events that occur in the world. Because of the large number of press articles, it is necessary to employ a large number of people to read press articles and direct readers to read useful and good press articles, which they are called critics. Although many critics are employed, they cannot read all the published press articles because it is not possible for critics to be familiar with all published newspaper articles. This will be the reason for the neglect of many useful articles, which may also be excellent.

Historically, the culture of interacting with news articles has been limited because readers who want to interact send interactive messages to the editor or by voice by telephone. As a result, it was difficult to form the so-called public forums because editors could not publish all readers' messages because of the maximum number of newspapers allowed. Nowadays, the masses of media have the ability to comment and discuss articles published on the Internet through public comment forums, which are ubiquitous through the Internet. It is now possible for the online newspaper community to participate in these public forums by commenting on news articles as they interact [1].

The availability of articles on the internet that have become easy to access by the society led to increase comment and interaction directly with the articles. The articles are commented on by people who differ in terms of age, cultural achievement, their experience with the article, and their goals of commenting [2].

The electronic press has many features that distinguish it from the rest of the traditional and new media, where it combines the readers and the audiovisual in one way, which means multiple patterns and processes of receiving content by the receiver (user) [3].

The use of electronic newspapers is a combination of reading the written texts and viewing images and videos, listening to audio files and interaction with multiple media content by the user, whether to send messages to the contactor or the newspaper management and to comment on the content or read the comments of other users [4].

The objectives of this paper is to use an automated scale instead of critics to evaluate articles by proposing a complete system to build the corpus that contains Arabic words that have positive characteristic and negative characteristic, and this system is useful for arranging a large block of data that is used to automatically evaluate articles. The main aim from the proposed system is to use a measure that is an alternative to critics in order to guide readers to read good and useful articles.

2. Theoretical Background

2.1. Literature Survey

There are many papers that show interest in electronic newspapers and readers' comments on articles as well as interest in the construction of corpus, and here we will show some of the previous work, which will be briefly explained:

In 2010, Ashraf AbdelRaouf et al. [5] described the construction and provides a comprehensive study and analysis of a multi-modal Arabic corpus (MMAC) that is suitable for use in both OCR development and linguistics. MMAC currently contains six million Arabic words and, unlike previous corpora, also includes connected segments or pieces of Arabic words (PAWs) as well as naked pieces of Arabic words (NPAWs) and naked words (NWords); PAWs and Words without diacritical marks. Multi-modal data is generated from both text, gathered from a wide variety of sources, and images of existing documents. Text-based data is complemented by a set of artificially generated images showing each of the Words, NWords, PAWs and NPAWs involved. Applications are provided to generate a natural-looking degradation to the generated images. A statistical analysis and verification of the dataset has been carried out and is presented .The MMAC corpus was designed to meet the needs of users of traditional linguistic corpora, but at the same time to be beneficial to OCR applications developers. It can be used in testing and training each phase of an Arabic OCR application.

In 2013 Mahmoud El-Haj & Rim Koulali [6] generated an Arabic multipurpose corpus, which have been called KALIMAT1 (Arabic transliteration of "WORDS"). The automatically created corpus could benefit researchers working on different Arabic NLP areas. In their work on Arabic they developed, enhanced and tested many Arabic NLP tools. They tuned these tools to provide high quality results. The tools include auto-summarisers, Part of Speech Tagger, Morphological Analyzer and Named Entity Recognition (NER). They have been ran these tools using the same

document collection. They provided the output corpus freely for researchers to evaluate their work and to run experiments for different Arabic NLP purposes using one corpus.

In 2018 Chan Woo Kim et al [7], proposed a set of methods for identifying and operationalizing controversial news items. Based on an analysis of online readers' comments posted in response to political news, they developed what they term as "controversy indicator". They calculated a controversy score using the total number of reader comments and the proportion of "upvotes" (indicating approval) and "downvotes" (indicating disapproval) a given news article elicits. Based on an analysis of the political news articles published on the Naver Web portal during the 2017 presidential election in South Korea, and readers' responses to them. They compared a group of articles that attracted the largest share of reader comments with a group of articles characterized by strong disagreements between readers. The former is denoted the "high positive indicator" group and the latter the "high negative indicator" group. While a positive indicator reveals the intensity of people's attention, a negative indicator reflects conflict and division among people. They found the potential usefulness of the controversy indicator in understanding the contemporary news environment, which is becoming increasingly divisive and polluted with disinformation.

2.2. Readers Interact with Articles through Comments [8]

The commentary is used by providing an online response to the opportunity for readers to interact with each other or express their consent or disagreement with the content of the article or the comments of others. Commentators also work to correct what they believe to be false information in other comments, provide support, share content, and provide facts and links to websites that are relevant to the article. Through public comments, a dialogue is created among commentators so that the ideas are discussed and negotiated. In fact, comments cannot be considered to represent the views of all people, but because of the large number of comments that are available on certain articles, it can show the views of large citizens. Comments and responses can be considered as a measure of immediate and honest spontaneous public opinion. Comments have an important role in shaping the reader's attitudes.

2.3. Concept of Corpus [9]

Corpus is a set of naturally occurring language statements that are either written texts or a copy of recorded speech. Corpus is a starting point for studying language or checking assumptions about the language. There are three main stages for building the corpus:

1. The preparatory step: this is about the work carried out before the corpus collection. In this step, key questions must be answered, such as: Why do we need a corpus? What properties should such a corpus have? How can we collect this corpus?
2. The collection and the annotation of the corpus: this step covers the work necessary to construct the corpus in such a way that the objectives fixed in the preceding step can be reached.
3. The use of the corpus: this step is about the statistical analysis and/or the linguistic analysis of the contents of the corpus. This step can bring some insights into the studied linguistic subject. For example, you can try to calculate the number of syntactic constructions by knowing the thematic context or the type of text (medical text, journalistic text, etc.).

2.4. The Concept of Natural Language Processing (NLP) [10]

NLP is an interdisciplinary field that combines computational linguistics, computing science, cognitive science, and artificial intelligence. From a scientific perspective, NLP aims to model the cognitive mechanisms underlying the understanding and production of human languages. From an engineering perspective, NLP is concerned with how to develop novel practical applications to facilitate the interactions between computers and human languages. There are several Procedures of NLP that are used to work with Raw Text as the following:

1. Tokenization and Stop Words: The raw text data usually contains two parts, one of which is useful for extraction of properties and the other is not useful, so it causes noise when conducting the extraction of properties, and therefore some steps must be followed to get rid of the unhelpful part. When text is read by a computer, it is read as a single string, even the punctuation marks in the text, so a way must be found to separate individual text words into individual strings. This leads to the concept of tokenization, which is simply the process of separating the whole chain into several individual tokens, which are represented as a word or character. After we get a series of tokens, we then remove some of the tokens that cause noise affecting the process of extracting the properties because they are

abundant in the language, which is called stop words. The concept of getting rid of stop words is usually called the concept of removing stop words.

2. **The Bag-of-Words Model (BoW):** It is one of the easiest procedures used to extract properties. It is named because it simply aims to find the number of times a particular word is found within a block of text. The way this model works is to tokenization the raw text into tokens (words) then finding the count (frequency) for each token (word).

2.5 Statistical Method in Evaluation of Press Articles

The most famous methods of statistics are measures of central tendency or location. Location statistics give an indication of how big or how small the data set is. The most commonly used statistics are the arithmetic mean. Arithmetic Mean is the most popular and well known measure of central tendency. The mean is equal to the sum of all the values in the data set divided by the number of values in the data set, so the arithmetic mean (M) of observations $(x_{1……} x_n$) as shown in Eq. (1) [11] :

$$M = \frac{1}{n}(x_1 + x_2 + \cdots + x_n) = \frac{\sum_{i=1}^{n} xi}{n} \qquad (1)$$

Where (n) is the number of the elements.

3 The Proposed System

The proposed system consists of two systems namely the Building Corpus and the Applying Statistical Technique stage. The aim of this system is to evaluate the journalist and his press articles based of characteristics of comments on press articles.

3.1 Building Corpus stage

The corpus system is proposed to arrange a large block of words and easy handling of words. This system is very important because the results of the proposed system depend entirely on it because through this system was determined the characteristics of words in terms of positive and negative. This stage was started by collecting comments (dataset) on several press articles used in the building of the corpus. The number of comments collected were about 200 comments that resulted from the readers' commentary on 12 press articles. The source of these press articles is the website of Al-Ahram newspaper from the political scene section during 2014. Dataset comments that was collected will be read once this entered into the proposed system, then all comments are divided into a series of words to get a list containing the words of all comments. Then comparing the list of Arabic stopwords with the list of words then removing the similar words. The frequency is then calculated for each word in the list of words of comment as shown in Table 1. These principles can be applied using an algorithm (3.1). In fact, algorithm (3.1) is the BOW algorithm, but a step for removing Arabic stopwords was added to the algorithm.

Algorithm (3.1): The Bag-of-Words Model (BOW)
Inputs: SQLtable(comments), List of Arabic stopwords
Output: Frequency of Words

Begin
Step 1 : Foreach word in Dataset
          Foreach word in list of Arabic stopwords
            If word = Arabic stopword then remove the word from list of words
             Find the frequency for each word in list of words
          Dictionary(word, frequency)= frequency of words
     Next
  Next
End

Table 1. Samples of Words and their Frequencies

Word	Frequency	Word	Frequency	Word	Frequency
قتل	7	السرقة	1	الحميم	1
ذبح	1	اللصوصيه	٢	انصارهم	2
للقتل	1	اباحها	2	لانصارهم	1
بقتل	2	الحبيبه	2	انصارها	1
الحارات	1	بصديقه	2	الانصار	2
القتل	1	مناطق	1	رهط	٢
اغتيالهم	1	العزيز	4	بمنطقه	1
الاغتيالات	2	محبوب	3	منطقة	1

After obtained Table1, based on expert experience, the clustering process was applied to Table 1.The clustering process start by taking a word from Table 1 and comparing it with the remainder of the words. Each word that is similar in meaning will be taken as a member of the cluster. This cluster is called synonyms of the selected word. Because the Arabic stemming is not used, there are words that contain the Prefixes and/or the Suffixes and that is why such words were made members of the synonyms. Frequencies are also accumulated for the selected word and all words within the synonyms. This process is continued until Table 2 is obtained.

Table 2. Sample of word and its synonyms

Word	Synonyms	Frequency
قتل	الاغتيالات، اغتيالهم، القتل، بقتل، للقتل، ذبح	15
السرقة	اللصوصيه، النشالين	3
الحبيبه	الحميم، محبوب، العزيز، عزيزي، اصدقاءه، بصديقه، الود، ود	12
انصارهم	لانصارهم، انصارها، الانصار، رهط	8
منطقة	بمنطقه، مناطق، الحارات	4
اباحها		٢

After the completion of the clustering process, Table 2 contains some words with their synonyms that have a small frequency, so they must be normalized. This accomplished using the algorithm (3.2).

Algorithm (3.2): Normalization process
Inputs: Corpus
Output: Improved corpus

Begin
Step 1: For each row in the corpus
If frequency < = ٣ then delete row from the corpus
 End If
Next
End

Algorithm (3.2) checks the frequency of each row in Table 2. If the frequency is smaller or equal to the threshold, it will delete the row from Table 2.

Table 3. Sample of Normalization Process

Word	Synonyms	Frequency
قتل	الاغتيالات، اغتيالهم، القتل، بقتل، للقتل، ذبح	15
السرقة	اللصوصيه، النشالين	3
الحبيبه	الحميم، محبوب، العزيز، عزيزي، اصدقاءه، بصديقه، الود، ود	12
انصارهم	لانصارهم، انصارها، الانصار، رهط	8
منطقة	بمنطقه، مناطق، الحارات	4

Each word and its synonyms in Table 3 are processed by the expert, where it is based on expert opinion the positive word and its synonyms are given the label "1", while the negative word and its synonyms are given the label "-1". Neutral word and its synonyms are given the label "0". Then positive word and its cluster or negative word and its cluster are given weight within the range of 1 to 5 where it is also based on expert opinion and based on the strength

of the influence of the positive word, or one of its synonyms, and the negative word, or one of its attributes, on the evaluation of the article. The process of labelling and weighting is shown in Table 4.

Table 4. Weighting and Labelling Process

Word	Synonyms	Weight	Type
قتل	الاغتيالات، اغتيالهم، القتل، لقتل، بقتل، للقتل، ذبح	٤	-1
السرقة	اللصوصيه، النشالين	٢	-1
الحبيبه	الحميم، محبوب، العزيز، عزيزة، عزيزي، اصدقاءه، بصديقه، الود، ود	١	١
انصارهم	لانصارهم، انصارها، الانصار، رهط، أتباع	٣	١
منطقة	بمنطقه، مناطق، الحارات	0	.

3.2 Applying Statistical Technique stage

   After building the corpus by building the corpus system, the corpus will be entered to the next step that applies arithmetic mean technique, as in algorithm (3.3).

Algorithm (3.3): Proposed Statistical Technique
Input: SQLtable, c: Corpus, cm: Comments, st: List of Arabic stop words
Output: Evaluation of comments
Begin **Step 1 :** Input database comments to be evaluated 　　**Foreach** comment in set of comments 　　Partition the comment into list of words then Remove Arabic stopwords from a list of comment words **Step 2:** input the corpus (Table 4) 　　**Foreach** record in the corpus 　　generate list that contains word and its synonyms **Step 3 :** Generate a lists of positive and negative words extracted from the comment **Step 4 :** Find the frequency for positive words and negative words **Step 5 :** Generate list of weights for positive words and list of weights for negative words by applying the equation 　$w_y = F * w_x$ **Step 6:** Calculate arithmetic mean statistic for each list of weights. 　$M = \frac{1}{n}(x_1 + x_2 + \cdots + x_n) = \frac{\sum_{i=1}^{n} xi}{n}$ **Step 7 :** Compare of statistical results between the list of weights of positive words and the list of weights of negative words 　**Next** 　**Next** 　**End**

The procedures of the proposed system can be explained by applying algorithm (3.3) to find the arithmetic mean statistic in order to evaluate the comments that lead to evaluate the articles which can be clarified as in the following steps:

1. Partition the comment into a list of single words. Then removing Arabic stopwords this is accomplished by comparing the list of Arabic stopwords with the list of comment words, then removing similar words from the list of comment words. The resulting word list after removing the stop words will be used in step (4). This step is repeated for each comment.
2. Read a row from Table 4 and represent it as a list containing the word and its synonyms. This step is repeated for each row from Table 4.
3. Generate the lists for positive and negative words that extract from the comment by comparing the resulting list from step (2) with the resulting list from step (3). If there is a similarity and the similar word is given label (1) in Table 4 then word is added to list of positive words. If there is a similarity and the similar word is given label (-1) in Table 4 then word is added to list of negative words.
4. Find the frequency for each word in the lists of positive and negative words using algorithm (3.1).
5. Generate a list of positive words weights and a list of negative words weights by comparing the resulting each lists from step (٤) with the result list from step (2). If there is a similarity and the similar word is given label (1) Table 4, then the weight is added to list of positive words. If there is a similarity and the similar word is given label (-1) Table 4,

then the weight is added to list of negative words. Then multiplying the frequency from step (4) by the lists of positive weights and negative weights by applying Eq. (2).

$$w_y = F * w_x \hspace{5cm} (2)$$

Where F the frequency of positive or negative words, w_x is the weight for the word

6. Calculate arithmetic mean statistic using Eq1 for each list of weights, in which the statistical result for positive and negative lists of weights are denoted by the symbol M_1 and M_2, respectively.
7. The evaluation of the comment is based on the comparison of results M_1 and M_2 results as well as on the thresholds within which the results are used where three thresholds were used which are symbolized by th_1, th_2 and th_3, respectively. The evaluation of the comment is as follows:

- In case the values M_1 and M_2 are within the th_1. If $M_1 > M_2$ then the evaluation of the comment is positive low elseif $M_2 > M_1$ then the evaluation of comment is negative low elseif $M_1 = M_2$ then the evaluation of comment is normal.
- In case the values M_1 and M_2 are within the th_2. If $M_1 > M_2$ then the evaluation of the comment is positive medial elseif $M_2 > M_1$ then the evaluation of comment is negative medial elseif $M_1 = M_2$ then the evaluation of comment is normal.
- In case the values M_1 and M_2 are within the th_3. If $M_1 > M_2$ then the evaluation of the comment is positive high elseif $M_2 > M_1$ then the evaluation of comment is negative high elseif $M_1 = M_2$ then the evaluation of comment is normal.

3.3. Evaluation Comments by the Experts

The questionnaire form is one of the most appropriate research tools that enable the widest possible access to information and data and allows saving time, effort and expenses, and achieving results that can be disseminated and invoked in the questionnaire through the form [12]. Each expert was asked to extract positive words and negative words and then give weight to each word whether positive or negative. Then the expert was asked to take the plural of the weights of the positive words and the weights of the negative words and then compare the results of the plural to evaluate the comment. Ten copies of each comment were made from the comments database and distributed to experts. After completing the questionnaire, the average of the plural results was taken to weigh the positive words and the weights of the negative words.

4. Results and Discussion

4.1 Database

The used database consist of two tables. The first table consist ten press articles for "اسامة ابو دومة" journalist. The second table consist of comments on press articles. It was used "one to many" relationship to connect the two tables. All of these press articles have a unique ID to distinguish them from each other and the comments take the ID based on ID of press articles. These press articles have been compiled from Al-Ahram newspaper from the "المشهد السياسي" section and these press articles date back to the year ٢٠١٤.

4.2 The proposed System Results

4.2.1 Corpus System Stage Results

Unlike other languages, the difficulty of the Arabic language is the lack of dealing with the Arabic language and the lack of corpus and corpus analysis tools, which led researchers to build their own corpus to suit their needs. The corpus was built to automated analysis of comments on press articles, which contain 429 clusters and 7300 single words.

4.2.2. The Automated Analysis of Press Articles

The press article is evaluated by collecting the comments of the press article as one comment and apply the proposed system on that collection of comments on the press article. If the result of the evaluation is "normal", "positive", "positive medial" or "positive high". The press article is worth reading by readers, but if the result of the evaluation is "normal", "negative", "negative medial" or "negative high", the press article is not worth reading.

Table 5 shown the results of the evaluation each press article and Table 6 shown the range of three threshold that that specific how strong a negative or positive press article.

Table 5. The Result for Each Press Article

Press Article ID	Evaluation of Press Articles	M_1	M_2
A1	Negative high	3	6
A2	Negative medial	2	3
A3	Negative medial	4	5
A4	Positive medial	4	3
A5	Normal	4	4
A6	Positive medial	4	3
A7	Positive medial	4	3
A8	Positive high	5	3
A9	Negative medial	2	4
A10	Negative high	2	6

It was noted in the Table 5, that the press articles A4, A5, A6, A7 and A8 were evaluated positively. Regardless of the strength of the positive characteristic, these articles were considered good and as a result readers were directed to these press articles. On the other hand, the press articles A1, A2, A3, A9 and A10 were evaluated negatively. Regardless of the strength of the negative characteristic, these articles were considered bad and as a result readers were directed not to read these press articles.

Table 6. Range Values of Thresholds

Threshold ID	Threshold Range	Evaluation of Press Article
TH1	>=2	"Positive" or "Negative"
TH2	>=3 and <=5	"Positive Medial" or "Negative Medial"
TH3	<=6	"Positive High" or Negative High"

4.3 Evaluation Comments by Experts Results

In section (3.3), it is explained how the comments were evaluated by experts, which was done using a questionnaire form. The results shown in Table 7 for each sample comment. These comments are the same as those evaluated by the proposed system.

Table 7 the Results for Each Press Article

Press Article ID	Evaluation of Press Articles	M_1	M_2
A1	Negative medial	3	5
A2	Negative medial	2	3
A3	Negative	4	٢
A4	Positive medial	4	3
A5	Normal	4	4
A6	Positive medial	4	3
A7	Positive medial	4	3
A8	Positive high	5	3
A9	Negative medial	2	4
A10	Negative high	2	6

4.3 Evaluation of the Proposed System

To show the effectiveness and performance of the proposed system, there are four measurements of evaluation: precision, recall, f-measure and accuracy, which have two types of corrects and two types of errors [13-15]:

1. **True positive (TP):** it's the number of words that have been given the label (1) by the expert as system.
2. **True negative (TN):** it's the number of words that have been given the label (-1) by the expert as system.
3. **False positive (FP):** it's the number of words given by the system is (1) while the expert has given it (-1).
4. **False negative (FN):** it's the number of words given by the system is (-1) while the expert has given it (1).
   These measurements have been explained as the following:

- **Precision:** it is the percentage of number of true positive words divided by the total number of positive words. This is calculated by the following equation:

$$\textbf{Precision} = \frac{TP}{(TP+FP)} \qquad\qquad (3)$$

- **Recall:** it is the percentage of the number of true positive words divided by the summation of number true positive and false negative words. This is calculated by the following equation:

$$\text{Recall} = \frac{TP}{(TP+FN)} \tag{4}$$

- **F-measure**: it is calculates the proportion of the multiplied double of the result of the first method (Precision) divided by the result of the second method (Recall) as a summation of the result of these methods: calculated by the following equation:

$$\text{F-measure} = \frac{2*Precision*Recall}{Pecision+Recall} \tag{5}$$

- **Accuracy**: it is the percentage of the summation the number of true positive and true negative words divided by the total number of words. This is calculated by the following equation:

$$\textbf{Accuracy} = \frac{(TP+TN)}{(TP+TN+FP+FN)} \tag{6}$$

 Several measures were used in evaluating the effectiveness of the proposed system's performance, because if one measure was used, it could be biased for the proposed system and an accurate evaluation could not be obtained. In the case of using more than one scale, this has given a more accurate evaluation.

Table 8: The Proposed System Evaluation Results

ID	Precision	Recall	F-measure	Accuracy	TP	TN	FP	FN
A1	0.85	0.81	0.83	0.87	22	40	4	5
A2	0.91	1	0.95	0.95	20	19	2	0
A3	0.89	0.8	0.84	0.85	24	29	3	6
A4	0.92	0.96	0.94	0.93	23	20	2	1
A5	1	1	1	1	54	33	0	0
A6	0.94	0.97	0.95	0.93	32	11	2	1
A7	0.95	1	0.97	0.97	20	14	1	0
A8	1	0.95	0.97	0.97	39	16	0	2
A9	1	0.95	0.97	0.97	20	28	0	5
A10	0.92	1	0.96	0.97	23	36	2	0

Table (4.8) illustrate precision, recall, f-measure and accuracy rustles for each press article. It's noted, that the evaluation of the A_1 by experts is "negative medial" instead of "negative high". Although the expert evaluation and the proposed system of the article is negative, the difference is in the strength of the negative as the result of the evaluation of the proposed system is negative as observed in the Table 5 and the result of the expert evaluation is negative as observed in the Table 7. This resulted in a percentage error in the evaluation of the article as the results. So the error ratio was about (15), (19), (17) and (13) for precision, recall, f-measure and accuracy, respectively. The A_3 by experts is "negative" instead of "negative medial". Although the expert evaluation and the proposed system of the article is negative, the difference is in the strength of the negative as the result of the evaluation of the proposed system is negative as observed in the Table 5 and the result of the expert evaluation is negative as observed in the Table 7. This resulted in a percentage error in the evaluation of the article as the results. So the error ratio was about (١١), (٢٠), (١٦) and (١٧) for precision, recall, f-measure and accuracy, respectively. As for the rest of the articles, although there is an error rate, it is so low that there is no difference in the results between the evaluation of the proposed system and the evaluation of experts.
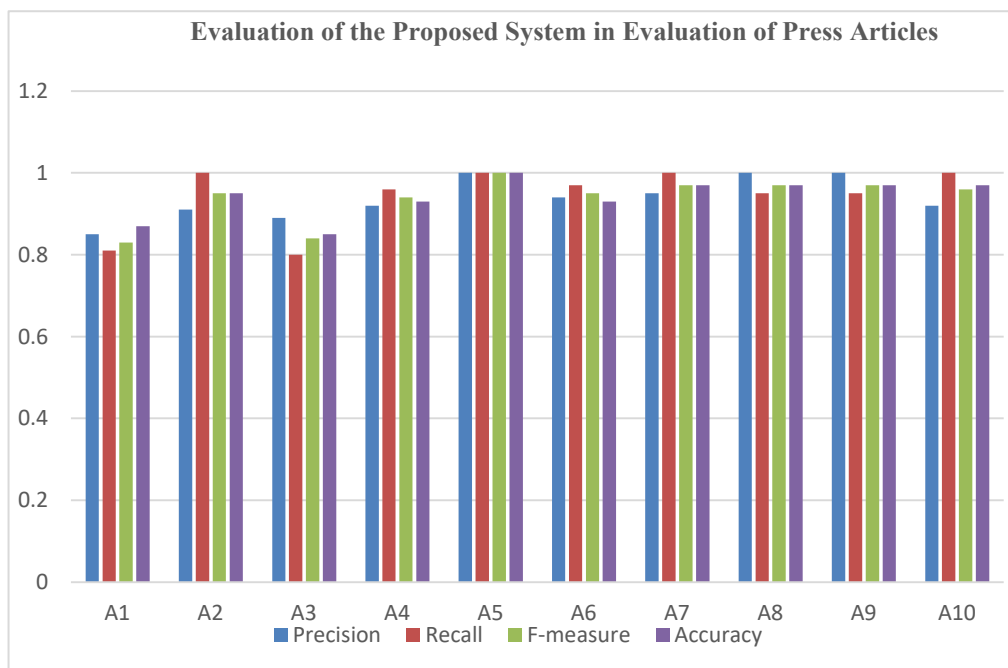
Fig. 1. The Proposed System Results Evaluation Chart

In order to evaluate the writer, all comments were collected for all the database press articles of the writer" حازم ابو دومة" as if they were one comment and then the proposed system applied to that one comment, as shown in Table 9 the evaluation of the writer is "negative", so this writer was considered bad and as a result readers were directed not to read the press articles for this writer

Table 9. Evaluation of "حازم ابو دومة"

Evaluation of Writer	M_1	M_2
Negative	5	7

As shown Table 9 the author's assessment is negative and therefore he did not recommend directing readers to read his press articles.

The evaluation of the proposed system in evaluate the writer "حازم ابو دومة" based on comments of press articles of him was shown in table Table 10

Table 10. Proposed system in evaluate the writer "حازم ابو دومة"

Precision	Recall	F-measure	Accuracy	TP	TN	FP	FN
0.95	0.95	0.95	0.94	277	246	16	16

As shown in Table 10 the evaluation results were excellent that was indicated the evaluation of the proposed system in evaluate the writer "حازم ابو دومة" was convergent from the opinion of experts.

The recall, precision, accuracy and the F-measure of the proclitics, as well as of the enclitics were evaluated. The total number of the measured cliticsis 526. The aim of the experiment is to measure the ability of the analyzer in tokenizing and predicting the proclitic particles such as conjunctions, vocative particles ,etc, and the enclitic pronouns such as possessive pronouns. The recall, precision, accuracy and the F-score of the previous features were calculated and the results were displayed. A model to calculate the four values of the confusion matrix TP/ FP/ TN/ FN was built, the model accepts manually designed Gold Standard and data sets, calculate the values of the desired morphological features then it measures the following matrixes: recall, precision, F-measure, accuracy . The mean average of the

final results was calculated for all the values to provide detailed information about the accuracy and the ability of the analyzer to predict the correct analysis (Fig.11 and Table2) [13].

Table11 Recall and Precision percentage

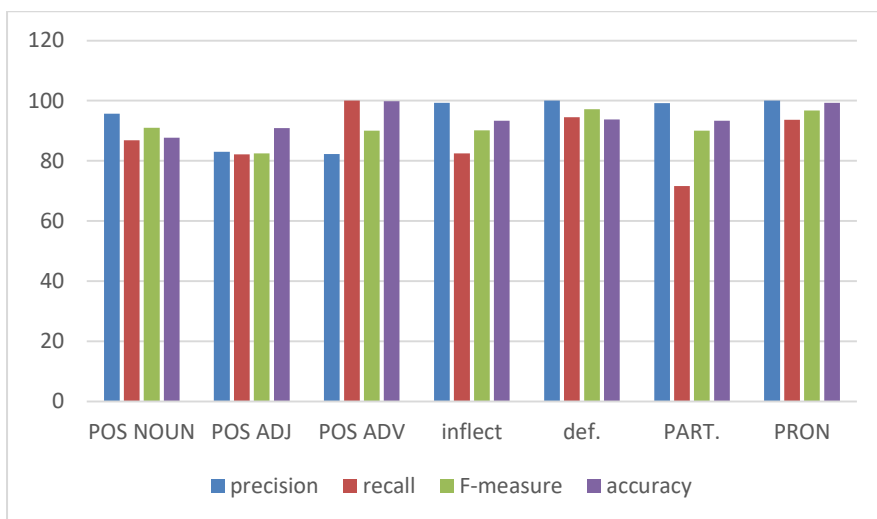| Feature	Precision(%)	Recall (%)	F-measure (%)	Accuracy (%)
Part of speech of noun	95.7	86.9	91	87.7
Part of speech of adjective	83	82.2	82.5	90.9
Part of speech of adverbs	82.3	100	90	99.8
Inflections	99.3	82.5	90.1	93.3
Definiteness	100	94.5	97.2	93.8
Proclitic particles	99.2	71.6	90	93.3
Enclitic pronoun	100	93.7	96.7	99.3

Fig.2 Recall and precision percentages of the morphological features

5. Conclusion and Future Works

This paper presents the implementation of the proposed system that was designed under the goal of using an automated scale to evaluate comments on press articles. Form the experimental results, some main conclusions are derived and presented in this section as follow:

- Its notes, that the performance of the proposed system on the comments database was very good and this indicates the extent of convergence between the performance of the proposed system and that of experts in evaluating comments. Thus, the proposed system is able to evaluate comments very efficiently.
- The performance of the proposed system in the evaluation of comments depends not only on the number of positive and negative words derived from the comment but also on the weight of words, where it is possible that the number of negative words is more than the number of positive words and the result of the evaluation of the comment is positive, because the weight of positive words is more and this indicates that the impact of positive words in the evaluation has more impact than negative words.
- The proposed system was applied to comments of different lengths and it was noted that the longer the comment the longer the performance of the proposed system which is more accurate in the evaluation and the reason is that it is possible to draw one positive word from the comment in the case of short comments and its strong impact on the comment .In the case of long comments, the more positive words appear, the less the impact.
- It was noted that corpus is not only limited to data collection but also has an active role in data analysis as it has been instrumental in analyzing comments and then categorizing them into positive and negative comments.
  For future works, it would be interesting to use stemmer for Arabic language in order to deal only with the roots of words instead of wasting time compiling words that contain additions and are considered synonyms. The use of artificial intelligence techniques in the evaluation of comments on press articles and using additional statistical techniques and comparing their results would be important.

References

1. L. Mathe and M. Caldwell," A Content Analysis of Readers' Comments on Political and Economic News Articles from Zimbabwe Online Media", Journal of Communication, 8(1), (2017), 56-68.
2. N. Diakopoulos and M. Naaman, "Towards Quality Discourse in Online News Comments". Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work. Rutgers, (2011), 133-142.
3. F.S Faiaz and H.K Milh, "News trends in the 2010 Iraqi Council of Representatives elections in the electronic press "(available in Arabic), Journal of Adab al-Farahidi, (24), (2017), 133-142.
4. K.H Hasan, "Patterns of use of electronic newspapers by the Iraqi public" (available in Arabic), Journal of Medad Arts, 1(16), (2019), 759-784.
5. A. AbdelRaouf, C.A Higgins, T. Pridmore and M. Khalil, "Building a Multi- Modal Arabic Corpus (MMAC)", International Journal on Document Analysis and Recognition, 13(4), (2010), 285-302.
6. M. El-Haj and R. Koulali, "Kalimat a Multipurpose Arabic Corpus", Second Workshop on Arabic Corpus Linguistics (WACL-2) , (2013), 22-25
7. C.W. Kim, H. Park, Y. Lee, H.W. Park, "Towards the Operationalization of Controversial News: A Study of Online News Articles and Reader Comments During the 2017 Presidential Election in South Korea". Journal of   Quality & Quantity, 53(3), (2018), 1097-1108.
8. N. Henrich and B. Holmes, "Web News Readers' Comments: Towards Developing A Ethodology for using On-Line Comments in Social Inquiry", Journal of Media and Communication Studies, 5(1), (2013), 1-4
9. M.Z. Kurdi, "Natural Language Processing and Computational Linguistics", UK and USA: ISTE Ltd and John & Sons, Inc, (2016).
10. L. Deng and Y. Liu, "Deep Learning in Natural Language Processing", USA and China: Springer Nature Singapore Pte Ltd, (2018).
11. P. Goos and D. Meintrup, "Statistics with JMP: Graphs, Descriptive Statistics, and Probability", India : John Wiley & Sons, (2015).
12. K. Soania, . (2011). "Attitudes of the student audience towards electronic journalism A field study on a sample of university students using the Internet" (available in Arabic) M.Sc. Thesis, University of Muhammd Kider, Sukra, Algeria, 2011.
13. K. Shaalan, A. Hassanien, and F. Tolba, "Intelligent Natural Language Processing: Trends and Applications", United Arab Emirates and Egypt: Springer International Publishing, (2018).
14. D. S. Al-Azzawy and F.M. Al-Rufaye. "Arabic Words Clustering by using K- Means Algorithm ".Annual Conference on New Trends in Information & Communications Technology Applications (NTICT), IEEE, (2017), 263-267.
15. J. Eisenstein, "Natural Language Processing", (2018).