

DOI: <http://dx.doi.org/10.21123/bsj.2022.19.2.0422>

Tourism Companies Assessment via Social Media Using Sentiment Analysis

Nadia F. AL-Bakri^{1}*

Janan Farag Yonan²
Ali Sami Abid³

Ahmed T. Sadiq³

¹ Department of Computer Science, College of Science, AL Nahrain University, Baghdad, Iraq.

² Minister Office, Ministry of Higher Education and Scientific Research, Baghdad, Iraq.

³ Department of Computer Science, University of Technology-Iraq, Baghdad, Iraq.

*Corresponding author: nadia.f.al-bakri@nahrainuniv.edu.iq jananfarag@gmail.com,

ahmed.t.sadiq@uotechnology.edu.iq

*ORCID ID: <https://orcid.org/0000-0001-6870-8572> , <https://orcid.org/0000-0002-7782-2091> , <https://orcid.org/0000-0002-4217-1321>

Received 28/9/2020, Accepted 2/3/2021, Published Online First 20/9/2021, Published 1/4/2022



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

Abstract:

In recent years, social media has been increasing widely and obviously as a media for users expressing their emotions and feelings through thousands of posts and comments related to tourism companies. As a consequence, it became difficult for tourists to read all the comments to determine whether these opinions are positive or negative to assess the success of a tourism company. In this paper, a modest model is proposed to assess e-tourism companies using Iraqi dialect reviews collected from Facebook. The reviews are analyzed using text mining techniques for sentiment classification. The generated sentiment words are classified into positive, negative and neutral comments by utilizing Rough Set Theory, Naïve Bayes and K-Nearest Neighbor methods. After experimental results, it was determined that out of 71 tested Iraqi tourism companies, 28% from these companies have very good assessment, 26% from these companies have good assessment, 31% from these companies have medium assessment, 4% from these companies have acceptance assessment and 11% from these companies have bad assessment. These results helped the companies to improve their work and programs responding sufficiently and quickly to customer demands.

Keywords: Facebook Data, K-NN, Naïve Bayes, Rough Set Theory, Sentiment Analysis, Text Mining.

Introduction:

Nowadays, online websites are suffering from lag in marketing their products due to lack of effective systems that analyze and trace customer assessments to their services; so some companies remain unknown despite their good quality of services ¹. Social media such as Facebook, Instagram and Twitter have taken a substantial part of publics and actions due to the huge and rapid advances in information technology. Many users are using communication sites not only to discover a new associate and share materials, but also to show their moods through variety of ways such as an electronic posting wall of comments. As such, Arab countries counted a huge number of social media users and became an appropriate media to endorse freedom of speech therefore; this media had become a rich resource for text mining and sentiment analysis methods ². Text Mining is used to detect undiscovered information on social media to obtain

useful information from different comments and reviews. Furthermore, web text mining uses data mining techniques to automatically extract and analyze information for knowledge discovery. Web data is typically unlabeled, distributed, heterogeneous, semi-structured, time varying, and high dimensional ^{3, 4}. Facebook is one of the most important advertising and promotion platforms for products, businesses and among the most important of these businesses is the tourism, travel and promotion for companies in this sector. Almost all tourism companies have a page on Facebook, through which they publish advertisements, promotion and propaganda for their travels, tourist trips and the services provided by these companies ⁵. Millions of people travel daily around the world for business, vacations, sightseeing and a large amount of money is spent on tickets, rooms, food, conveyance and enjoyment ⁶.

Two types of information flow in tourism business, the first flow is from the company to tourists such as tickets and different types of reservations and the other flow is the aggregate information from the tourists to the company. These two flows of information can be provided and analyzed by a suitable text mining algorithms to explore meaningful and vital decisions for the tourism provider. The essential part of information collecting manner is finding the emotions and thinking of other users. With the growing popularity of online review sites, new challenges are arisen to use information extraction techniques to detect and understand the sentiments of others⁷. Ultimately, sentiment analysis strategy checks assessments of users; for example, an individual tourist in correspondence with a tourist supports agent. It observes negotiations and assesses discussion to evaluate tempers and emotional states through a scoring mechanism, especially those associated with a business works⁸.

The main contributions of this paper are:

- 1-The main contribution is how to extract tourist's opinions that are available in tourism's company posts on Facebook and how it can be used to draw meaningful decisions about tourism companies indicating positive and negative sentiments.
- 2- This paper focuses on Arabic language with Iraqi vernacular text classification because it has affluent morphology. Most of researches deal with classical English and Arabic languages only.

In this paper, a sentiment analysis model for Tourist Company is suggested on Iraqi dialect Facebook posts to assess tourism and travel companies based on extracting sentiments from customer's comments on the social networking sites of these companies to discover useful decisions. The rest of this paper is organized as follows: In section 2, the related works on this field is subjected. In section 3, the general concepts of sentiment analysis are shown and processes are summed up. In section 4, the proposed model is presented. In section 5, the experiments on the suggested model are conducted. Last section is the conclusion of this work.

Related Works:

In what follows, some of the previous researches related to the techniques used for sentiment analysis are presented:

- 1- **In (9)**, the authors incorporated sentiment analysis methods to analyze reviews from travel blogs. Naïve Bayes, SVM and N-gram character based are sentiment classification methods used

for classifying reviews from travel blogs for 7 popular destinations in Europe and the US. Experimental results indicated that when a large number of reviews was analyzed all three methods approximately approached at least 80% of accuracies.

- 2- **In (7)**, sentiment analysis is performed on Arabic Facebook news pages. The proposed system consists of preprocessing, feature vector selections and classification methods using support vector machines (SVM), Naive Bayes (NB) and decision tree. 2400 comments have been collected and represented into a distinct record, and then grouped into 3 groups (supportive, attacking and neutral). After grouping, 800 are held for neutral comments, 800 for supportive comments, and 800 for attacking comments. The accuracy reached about 59.9% using Naive Bayes. The precision and recall are improved 10% with the decision tree and finally SVM gives the best for precision and recall which is about 73.4%.
- 3- **In (2)**, text mining techniques is explored for sentiment analysis. Tunisian users' statuses and behaviors are extracted from Facebook posts during the "Arabic Spring" to investigate the Tunisians' manners during the revolution in December and January at 2011. Classification techniques based on Support Vector Machine (SVM) and Naïve Bayes are utilized. A sentiment lexicon, based on the emoticons, interjections and acronyms are constructed from extracted statuses updates on Facebook. Comparative experiments are performed between two machine learning algorithms SVM and Naïve Bayes through a training model for sentiment classification. It was concluded that, using Naïve Bayes (NB), the accuracy reached high when the bigram feature was utilized. In contrast, the support vector machine (SVM) outperformed the Naïve Bayes (NB) when unigram feature is used.
- 4- **In (10)**, in this paper, a comparative study is applied between using Support Vector Machine (SVM), Naïve Bayesian (NB) and Multilayer Perceptron Neural Network (MLP-NN) classification methods on Arabic data sets which are Aljazeera news web site Saudi Press, Agency (SPA) and Alhayat. The experimental results were applied on 1400 Arabic documents belonging to different categories yielded a precision of about 0.778, 0.754, and 0.717 for SVM, NB and MPL-NN using 600 input layers respectively.
- 5- **In (11)**, this paper, an Arabic corpus is built consisting of Facebook (FB) posts written in

Dialectal Arabic (DA) with no grammar. The collections are labeled with five labels (positive, negative, dual, neutral, and spam). The words specifying opinions are used in a lexicon-based classifier.

- 6- In (12), an extraction model is proposed on a set of reviews from the public page ‘Opposing Views’ on Facebook using the software QSR NVivo 11 which is used to analyze unstructured data. The auto code feature in QSR NVivo 11 was utilized to analyze and tag the comments to positive or negative sentiments. Different Techniques were used for preprocessing such as Tokenization, Stemmed words and query augmentation for synonyms. After analyzing a specified number of comments, positive sentiments were about 29.6% and negative sentiments about 62.0%. The experimental results showed that the percentage of negative reviews is twice more than positive reviews.
- 7- In (3), a conceptual model based on text mining is proposed to classify texts into 4 emotional groups from the emotional dataset, they are: anger, fear, joy, and sadness. The Naïve Bayes optimized by particle swarm optimization (PSO) is utilized as a classifier. The preprocessing stages within model development were: document collection, transforming to lower

cases, tokenization, token filter (by length), stop words removal, stemming and vector creation. The experimental results yielded accuracy about 65.93% using Naïve Bayes without optimization, and 66.54% with PSO which showed an increase in the accuracy.

- 8- In (5), presents a lexicon-based sentiment analysis for common Iraqi vernacular. Three machine learning methods (KNN, Naive Bayes and Rough Set Theory) are used to classify Iraqi sentiment on Facebook. A dictionary of Iraqi keywords is built to include single and double word entries for both positive and negative sentiment. Rough Set Theory gives the best classification ratio compared with the other two methods. The classification ratio of rough set theory is better than KNN and Naïve Bayes because the upper and lower boundary in rough set provides a good tool to overcome the conflict problem in sentiment analysis. In Table 1, depicts the previous studies including the classifiers used, on what language the sentiment analysis is performed and their gaps. From these studies, it was identified that an Arabic language and Iraqi dialect based on sentiment analysis for tourism companies is not considered and it was proposed to develop the solution to the research problem.

Table 1. Previous Studies and their Specifications

Ref	Year	Aim	Language	Methodology	Gap
(9)	2009	Analyze reviews from travel blogs	English	Naïve Bayes SVM N-gram	Yahoo’s travel character based reviews since September 2007
(7)	2013	Comments are analyzed into Supportive, attacking and neutral	Arabic Facebook news pages	SVM Naive Bayes Decision Tree	Long comments not included Accuracy 59.9% Naive Bayes
(2)	2013	Status & behavior of Facebook users	Arabic	SVM Naive Bayes	No Vernacular Only bigram features
(10)	2016	a comparative study applied on 3 methods	Arabic data sets from Aljazeera news web site Saudi Press, Agency (SPA) and Alhayat	SVM Naïve Bayesian MLP-NN	No feature extraction and selection is applied
(11)	2017	built Arabic corpora consisting of Facebook (FB) posts	Dialectal Arabic (DA)	Lexicon-based classifier	Dialectal Arabic (DA) with no grammar
(12)	2018	Analyze unstructured data into positive or negative comments.	English from the public page ‘Opposing Views’ on Facebook	QSR NVivo 11 Software	NVivo does not attempt to classify a whole comment It classifies only words
(3)	2019	Classify text into 4 emotional groups: anger, fear, joy, and sadness	English data set WASSA-2017	Naïve Bayes	Feature selection using PSO
(5)	2020	Classify the comments, posts and pages in Facebook	Arabic language - Iraqi Vernacular	Rough set theory, KNN, Naïve Bayes	Some important semantic features are dropped

Sentiment Analysis:

Sentiment analysis (SA) is first presented by Liu, B¹³. It is also known as opinion mining which is the process of determining the polarity of opinions or reviews written by humans to rate products or services. SA can be done on a document level where the entire text is assessed to determine the opinion polarity by extracting features. Also it can be done on sentence level where the text is partitioned into sentences to be evaluated separately to determine the document polarity. Text polarity can be positive, negative or neutral. In tourism, information provided through a comment is either subjective (opinionated) which are based on personality feelings and decision making about events or objective (factual) which are based on information, evidences, and opinions. Sentiment analysis is utilized for evaluating written or vocalized languages to recognize whether speech is positive, negative or neutral and to what degree. Currently, there are many analysis tools able to deal with remarkable volumes of customer quibbles consistently⁸.

The need for Facebook has increased with the spread of COVID-19, forcing people to work from home without going to the workplace. From here, the strength of the research emerges through knowing the tendencies of the customers of the travel and tourism companies, and through these trends we can evaluate the company's. Tourist's comments and posts often reveal pleasure, prevention, displeasure, enjoyment. This sentiment information has a significant impact for tourism companies seeking to enhance customer managing and commercial productivity¹⁴. Facebook has particularly instigated long-term customer customization and rapid growth in terms of page size, comments and posts⁵.

Sentiment analysis includes a multi-step operations, they are: a) retrieval of data, b) extraction and selection the required data, c) pre-processing, d) feature extraction, e) subject detection, and f) applying data mining methods⁶. There are two fundamental methods for extracting sentiments. The first method is using lexical-based which involves manipulating a document polarity from the polarity of terms in the document. The second approach is using machine learning techniques which are classified into unsupervised and supervised methods. The supervised technique which is utilized in the proposed model involves classifier construction from labeled examples of texts through a supervised classification process. The advantage of machine learning techniques is creating trained models for trained contents to classify new data. The key of machine learning

methods used for sentiment analysis are Support Vector Machine (SVM) and Naïve Bayes as they were usually designed for binary-class classification tasks⁶⁻¹⁵. Rough set theory plays a big role in text classification and categorization within different areas¹⁶⁻¹⁹ therefore it is utilized in this paper as a classifier. Furthermore, Naïve Bayes Classifier (NBC)²⁰⁻²⁵ is also utilized in this paper as a probabilistic classifier, which uses Bayes' theorem as a decision rule with independent features.

The Proposed System:

In this section, the fundamental stages of the proposed model are presented. Sentiment analysis of the texts included in the comments is the fundamental operation in the proposed model because the main objective of the proposed model is to assess tourism and travel companies based on extracting sentiments from customer's comments on the social networking sites of these companies to discover useful decisions. The following are the main stages of the proposed model as shown in Fig. 1.

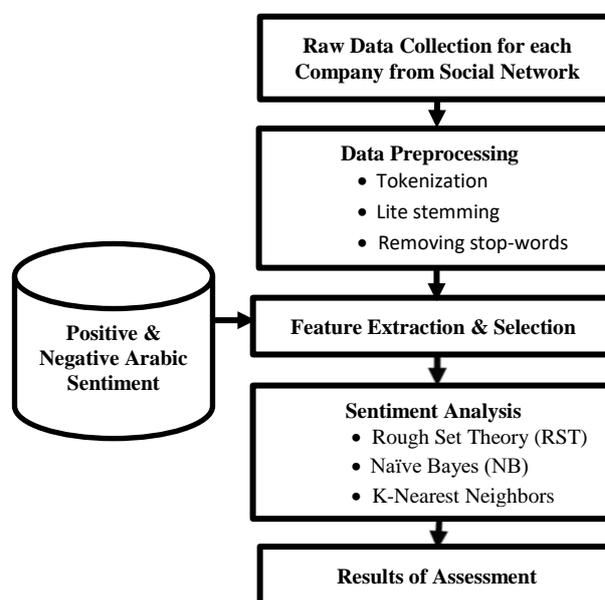


Figure 1. The Main Steps of the Proposed Model

- 1-Text Comments Collection Stage:** In the first stage, the customers as a source of information in social media leave feedback comments about visited tourism companies. So, these text comments are collected for each company separately to be further processed.
- 2-The Preprocessing Stage:** in this stage, a transformation is implemented to transform the extracted raw text comments into a suitable representation for the next stage. This involves three processes: tokenization process, stemming

process and removing stop-words process. Lite stemming has been used in the proposed model.

- 3- **Features Extraction and Selection Stage:** in this stage, sentiment words are selected (i.e. verbs and adjectives) and annotated as positive and negative sentiments. As Iraqi dialect includes all forms of spoken Arabic in day life, there are classical Arabic words within Iraqi dialect which are positive and negative words⁵. These words play an important role in the sentiment analysis and consequently on the assessment results effectiveness and quality. Two types of Arabic positive and negative words are considered in this stage. They are: single-word and double-word. It consists of 532 single-words (215 positive and 317 negative) and 419 double-words (233 positive and 186 negative).
- 4- **Sentiment Analysis Stage:** in this stage, three different types of machine learning algorithms are implemented in the sentiment analysis steps which are Rough Set Theory (RST), Naïve Bayes (NB) and K-Nearest Neighbors (KNN). Finally, depending on the sentiment analysis results which are either positive or negative, the assessment will be decided.

Results:

In this section, the proposed model is experimented with the collected information from Facebook. The dataset are collected from 71 Iraqi tourism companies in the Facebook, Table 2 shows the number of dataset details.

Table 2. Dataset Details

Number of Pages	Number of Posts	Number of comments	Number of Sentences
71	568	14200	33875

As mentioned in the above table, there are 71 company's pages in the Facebook, 8 posts from each company have been selected, 25 comments from each post have been selected, and the total sentences from all comments are 33875. The sentiment analysis took into account the classical Arabic language and Iraqi dialect words.

Three classification methods have been implemented for 14200 comments; these methods are Rough Set Theory, Naïve Bayes and K-Nearest Neighbor. Figure 2 illustrates the accuracy ratio of classification methods for sentiment analysis for the above dataset.

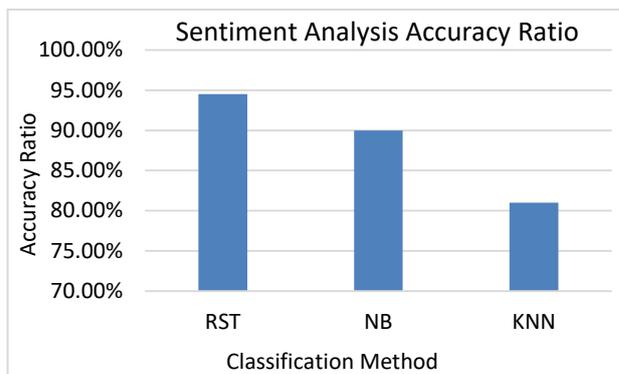


Figure 2. Sentiment Analysis Accuracy

The best one is Rough Set Theory (RST), (Table 3) illustrates the confusion matrix of RST classification method.

Table 3. Multiclass Confusion Matrix for 14200 Comments' Sentiment Classification

	Positive Predicted Class	Negative Predicted Class	Neutral Predicted Class	Total
Positive Actual Class	9642	160	58	9860
Negative Actual Class	542	3531	82	4155
Neutral Actual Class	3	7	175	185
				14200

In more details, below Tables 4, 5 and 6 show the confusion matrices for the positive, negative and neutral sentiment of the comments respectively.

Table 4. Confusion Matrix for Positive Sentiment

	Positive	Non-Positive	Total
Positive	9642	218	9860
Non-Positive	545	3795	4340
			14200

Table 5. Confusion Matrix for Negative Sentiment

	Negative	Non-Negative	Total
Negative	3531	624	4155
Non-Negative	167	9878	10045
			14200

Table 6. Confusion Matrix for Neutral Sentiment

	Neutral	Non-Neutral	Total
Neutral	175	10	185
Non-Neutral	140	13875	14015
			14200

The accuracy equation for binary classifier is calculated as follows¹³:

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \dots (1)$$

From above Eq. (1) and using data from Table 3, 4 and 5 it could be shown that the accuracy reached about 94.62% on positive comments, whereas reached about 94.42% and 98.9% on negative and neutral comments. Furthermore, the average accuracy¹³ which is the average of per-class effectiveness value of the classifier is calculated using the following Eq. (2). The resulted average accuracy is about 95.98%.

$$Average Accuracy = \frac{\sum_{i=1}^l \frac{TP_i + TN_i}{TP_i + FN_i + FP_i + TN_i}}{l} \dots (2)$$

Where l is total number of classes.

The average Error Rate¹³ which is the average of per-class classification error is computed using the following Eq. (3), which resulted about 3.9%.

$$Average Error Rate = \frac{\sum_{i=1}^l \frac{FP_i + FN_i}{TP_i + FN_i + FP_i + TN_i}}{l} \dots (3)$$

The following equation Eq. (4) shows the precision computation¹³ which computed the average per-class agreement of the human judgment with the classification model which yielded about 92.36%.

$$Precision = \frac{\sum_{i=1}^l \frac{TP_i}{TP_i + FP_i}}{l} \dots (4)$$

The following equation Eq. (5) shows the recall computation¹³ which computed the average per-class effectiveness of the model to identify the judgment of human which yielded about 81.8%.

$$Recall = \frac{\sum_{i=1}^l \frac{TP_i}{TP_i + FN_i}}{l} \dots (5)$$

The assessment of tourism companies depend on the sentiment of customer's comments in the Facebook social network. So the following rules illustrate the assessment of Tourism Company depending on sentiment analysis results:

If (Positive Sentiment $\geq 80\%$) then Assessment is Very Good

If (80% \geq Positive Sentiment \geq 70%) then Assessment is Good

If (70% \geq Positive Sentiment \geq 60%) and (20% \geq Negative Sentiment \geq 10%) then Assessment is Medium

If (60% \geq Positive Sentiment \geq 50%) and (30% \geq Negative Sentiment \geq 10%) then Assessment is Acceptance

Otherwise the Assessment is Bad

By applying the above rules on 71 Iraqi tourism companies, 28% from these companies have very good assessment, 26% from these companies have good assessment, 31% from these companies have medium assessment, 4% from these companies have acceptance assessment and 11% from these companies have bad assessment. Fig. 3 illustrates these results.

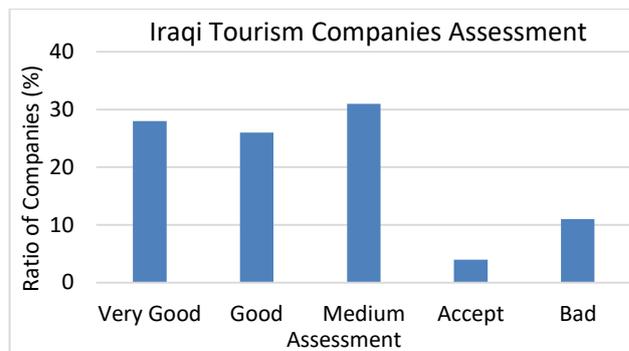


Figure 3. Iraqi Tourism Companies Assessment

Conclusion:

Sentiment analysis of Iraqi dialect reviews in Facebook has been proposed for the assessment of Iraqi tourism companies. Positive and negative Iraqi dialect sentiments are represented as single and double words (Bag-of-Words) to classify the sentiments. From the experiments, it was observed that the percentage results of our proposed model represent the real opinions of tourists in these tested companies. Utilizing standard Arabic language and Iraqi dialect sentiment analysis played a big role in our model. Rough set theory gave the best classification accuracy compared with Naïve Bayes and K-Nearest Neighbors, because this theory has shown a great benefit for categorized data. Also, it is concluded that instead of utilizing costly studies based on surveys, the sentiment analysis of customer opinions for tourisms made easier for companies to recognize their economical values and customers' opinions about their services, which in turns provides insight into enhanced decision making policies. The limitation of the proposed work was the absence to some semantic features; therefore it is better to consider these features as a future work.

Authors' declaration:

- Conflicts of Interest: None.
- We hereby confirm that all the Figures and Tables in the manuscript are mine ours. Besides,

the Figures and images, which are not mine ours, have been given the permission for re-publication attached with the manuscript.

- Ethical Clearance: The project was approved by the local ethical committee in AL Nahrain University.

Authors' contributions statement:

- Conception, design, drafting the MS, revision and proofreading: Nadia F. AL-Bakri.
- Analysis, interpretation: Janan Farag Yonan.
- Conception, design, acquisition of data, analysis, drafting the MS: Ahmed T. Sadiq.
- Acquisition of data, analysis: Ali Sami Abid.

References:

1. AL-Bakri N.F., Hashim S.H. A modified similarity measure for improving accuracy of user-based collaborative filtering. *IJS*. 2018; 59(2): 934-945.
2. Akaichi J, Dhouioui Z, Pérez MJ. Text mining facebook status updates for sentiment classification. *ICSTCC*. 2013 Oct 11: 640-645. IEEE.
3. Junianto E, Rachman R. Implementation of Text Mining Model to Emotions Detection on Social Media Comments Using Particle Swarm Optimization and Naive Bayes Classifier. *CITSM*. 2019 Nov 6; 7: 1-6. IEEE.
4. AL-Bakri N.F., Al-zubidi Azhar F., Alnajjar Aseel B., Qahtan Ehsan .multi label restaurant classification using support vector machine. *PEN*.2021; 9(2):774-783.
5. Ali AT, Ahmed T S. Lexicon-Based Sentiment Analysis for Iraqi Vernacular/Arabic Language Using KNN, Naive Bayes and Rough Set Theory. *JARDCS*.2020; 11(11).
6. Alaei AR, Becken S, Stantic B. Sentiment analysis in tourism: capitalizing on big data. *JTR*. 2019 Feb; 58(2):175-91.
7. Hamouda AE, El-taher FE. Sentiment analyzer for arabic comments system. *Int. J. ACSA*. 2013 Mar; 4(3).
8. Ali AT, Ahmed T.S. Twitter Sentiment Analysis and Events Prediction Using Lexicon Sentiment Analysis for Iraqi Vernacular. *JSST*. 2020;63(6).
9. Ye Q, Zhang Z, Law R. Sentiment classification of online reviews to travel destinations by supervised machine learning approaches. *ESWA*. 2009 Apr 1;36(3):6527-35.
10. Mohammad A.H, Alwada'n T, Al-Momani O. Arabic text categorization using support vector machine, Naïve Bayes and neural network. *GSTF (JoC)*. 2016 Aug 1;5(1):108.
11. Itani M, Roast C, Al-Khayatt S. Corpora for sentiment analysis of Arabic text in social media. *ICICS*. 2017 Apr 4 (pp. 64-69). IEEE.
12. Pudaruth S, Moheeputh S, Permessur N, Chamroo A. Sentiment Analysis from Facebook Comments using Automatic Coding in NVivo 11. *ADCAIJ*. 2018; 7(1): pp 41-48.
13. Liu B. Sentiment analysis and subjectivity. *Handbook of natural language processing*. 2010 Feb;2(2010):627-66.
14. Younis E.M. Sentiment analysis and text mining for social media microblogs using open source tools: an empirical study. *IJCA*. 2015 Jan 1; 112(5).
15. González-Rodríguez MR, Martínez-Torres MR, Toral SL. Monitoring travel-related information on social media through sentiment analysis. In 2014 IEEE/ACM (UCC) 2014 Dec 8 (pp. 636-641). IEEE.
16. Alsaïdi S.A, Sadiq A.T, Abdullah H.S. English poems categorization using text mining and rough set theory. *BEEI*. 2020 Aug 1; 9(4):1701-10.
17. Hasanin A. M, Ahmed T.S. Arabic Text Mining and Rough Set Theory For Decision Support System. *JACSTR*. 2019; 9(4).
18. Dhafar H. A, Ahmed T.S., Ayad R.A. A New framework for Automatic Extraction Polarity and Target of Articles. *ACRIT*. 2019; 1174. Springer.
19. Saif A. A., Ahmed T.S., Hasanen S.A. Wisdom Extraction from English Poems Using Tow-Level Feature Selection with Rough Set Theory. *IJAST*.2020; 29(4).
20. Abd DH, Sadiq AT, Abbas AR. Political Articles Categorization Based on Different Naïve Bayes Models. *ACRIT*. 2019 Sep 15: 286-301. Springer, Cham.
21. Al-zubidi A.F., Al-Bakri N.F., Hasoun R.K., Hashim S.H., Alrikabi H.T.S .Mobile Application to Detect Covid-19 Pandemic by Using Classification Techniques: Proposed System. *iJIM*. 2021; 15(16): 34–51.
22. AL-Bakri N.F., Hashim S.H .A Study on the Accuracy of Prediction in Recommendation System Based on Similarity Measures. *BSJ*. 2019; 16(1):263-269.
23. AL-Bakri N.F., Hassan S. A Proposed Model to Solve Cold Start Problem using Fuzzy User-Based Clustering. *SCCS2019 IEEE*: 121-125.8852624.
24. AL-Bakri N.F., Hashim S.H. Collaborative Filtering Recommendation Model Based on k-means Clustering. *ANJS*. 2019 Mar 1; 22(1):74-79.
25. Muhsen A.R., Jumaa G.G., AL-Bakri N.F., Sadiq A.T. Feature Selection Strategy for Network Intrusion Detection System (NIDS) Using Meerkat Clan Algorithm. *iJIM*. 2021; 15(16):158-171.

تقييم شركات السياحة عبر وسائل التواصل الاجتماعي باستخدام تحليل المشاعر

علي سامي³أحمد طارق صادق³جنان فرج يونان²نادية فاضل البكري¹¹ قسم علوم الحاسبات، كلية العلوم، جامعة النهرين، بغداد، العراق.² مكتب الوزير، وزارة التعليم العالي والبحث العلمي، بغداد، العراق.³ قسم علوم الحاسبات، الجامعة التكنولوجية، بغداد، العراق.

الخلاصة:

ازدادت وسائل التواصل الاجتماعي بشكل كبير وواضح لانها وسيلة إعلام للمستخدمين للتعبير عن مشاعرهم من خلال آلاف المنشورات والتعليقات حول شركات السياحة. وبالتالي ، يصعب على السائح قراءة جميع التعليقات لتحديد ما إذا كانت تلك الآراء إيجابية أم سلبية لتقييم نجاح الشركة. في هذه البحث، تم استخدام التنقيب عن النص لتصنيف المشاعر من خلال جمع مراجعات اللهجة العراقية حول شركات السياحة من الفيس بوك لتحليلها باستخدام تحليل المشاعر لتتبع المشاعر الموجودة في المنشورات والتعليقات. ثم تم تصنيفها إلى تعليق إيجابي أو سلبي أو محايد باستخدام K-Nearest Neighbor ، Naïve Bayes, Rough Set Theory. من بين 71 شركة سياحة عراقية وجدت أن 28% من هذه الشركات لديها تقييم جيد جدا ، و 26% من هذه الشركات لديها تقييم جيد ، و 31% من هذه الشركات لديها تقييم متوسط ، و 4% من هذه الشركات لديها تقييم مقبول و 11% من هذه الشركات لديها تقييم سيء. ساعدت النتائج التجريبية الشركات على تحسين عملها وبرامجها واستجابة كافية وسريعة لمتطلبات العملاء.

الكلمات المفتاحية: بيانات الفيسبوك، أقرب تجاور، نايف بايز، نظرية المجموعة التقريبية، تحليل الميول، تنقيب النصوص.