A Computational-Augmented Critical Discourse Analysis of Tweets on the Saudi General Entertainment Authority Activities

Waheed M. A. Altohami[1,2], & Abdulfattah Omar [1, 3]

[1] Department of English, College of Science & Humanities, Prince Sattam Bin Abdulaziz University, Alkharj, Saudi Arabia

[2] Department of Foreign Languages, Faculty of Education, Mansoura University, Egypt

[3] Department of English, Faculty of Arts, Port Said University, Egypt

Correspondence: Waheed M. A. Altohami, Department of English, College of Science & Humanities, Prince Sattam Bin Abdulaziz University, Alkharj, Saudi Arabia.

Received: October 28, 2022 Accepted: December 5, 2022 Online Published: December 12, 2022

doi:10.5430/wjel.v12n8p471 URL: https://doi.org/10.5430/wjel.v12n8p471

Abstract

This study used both computational tools in the form of a machine learning predictive model (Support Vector Machine) and a critical discourse analysis model (Van Dijk's ideological square model) (Van Dijk, 1993, 2008, 2009) to fulfill three objectives: (1) clustering the Saudis' Twitter-based opinions and sentiments regarding the entertaining and recreational activities run by the Saudi General Entertainment Authority (GEA); (2) offering empirical evidence on how computational linguistic methods could be implemented for offering a reliable conceptual framing of such opinionated big data; and (3) outlining the central themes generating ideologically motivated polarity in Saudi public opinion and the macrostrategies through which this polarity is textually instantiated and actualized. Toward fulfilling these objectives, we designed a purpose-built corpus of 9378 tweets based on five trending hashtags, covering the period between 2020 and 2022. Findings affirmed the efficacy of synergizing the Support Vector Machine model and the ideological square model in clustering and interpreting the target tweets. Based on the output discourse features and thematization of the tweets, two main groups with different ideologically motivated perspectives were identified. This ideological polarity was achieved through the use of two macrostrategies: positive self-presentation and negative other-presentation. These findings may prompt policymakers to reconsider current (mis)practices in order to achieve long-term sustainable development goals.

Keywords: critical discourse analysis, data mining, General Entertainment Authority, ideological square, opinion mining, Twitter, vector space clustering

1. Introduction

In an attempt to promote culture and entertainment in the Kingdom of Saudi Arabia, the Saudi General Entertainment Authority (GEA) and the General Commission for Audiovisual Media (GCAM) have been established as part of the Saudi Vision 2030. One of the main concerns of Vision 2030 is to have partnerships with international entertainment corporations and to increase the budget for recreational services. The GEA was established in 2016, and its first and current chair is Turki Al-Sheikh. It demonstrated a strong interest in developing the film industry and organizing activities and events that drew international celebrities as well as audiences from across the kingdom. All these steps were intended to divert Saudi Arabia's GDP away from the oil sector and offer job and training opportunities for Saudis. However, these events sparked hectic debates among Saudis. While some Saudis considered these entertaining activities a chance to promote tourism and the quality of citizens' lives, some other Saudis were against them, claiming that they contradicted the teachings of Islam and might spoil the Saudi youth. Many of these debates were carried out on social media platforms. A plethora of hashtags on Twitter were created either supporting or rejecting this step taken by the Saudi government. Saudi women, in particular, considered establishing these entertaining activities and events as part of women's rights. Conservatives saw them as mere imitations of the West. Saudi officials have never addressed such a controversy, despite the fact that numerous activities and events are planned on a daily basis.

Opinion mining is a technique used for the automatic extraction and identification of opinions, sentiments, evaluations, and attitudes regarding a particular issue or topic using computational tools and other natural language processing tools (Liu, 2012; Pang & Lee, 2008). The present study aims at conducting a computationally-augmented critical discourse analysis of the opinions of Saudis regarding the activities and events organized by the GEA over Twitter, the most popular microblogging platform (Pak & Paroubek, 2010), with more than 2.4 million active users in Saudi Arabia (see Figure 1).

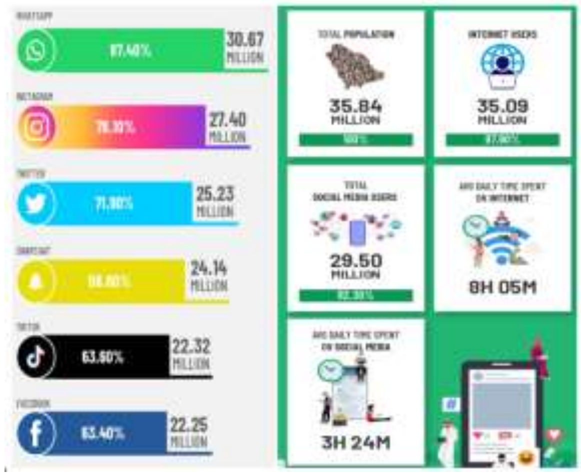

Figure 1. Key social media statistics in Saudi Arabia in 2022

This study focuses on Twitter in the Saudi context because it is the most popular social networking application among Saudis. Indeed, numerous studies have confirmed that Twitter has recently had a growing influence on Saudi public opinion on many of the issues that have sparked controversy in Saudi Arabia, such as the ban on women driving and the COVID-19 vaccines. Though creating an independent authority for promoting culture and entertainment in Saudi Arabia is an unprecedented event in a country with a long history of fundamentalism, little work has been done regarding the analysis of Saudis' reactions and responses toward the activities and events organized by the GEA. Therefore, this study sets out to provide an empirical answer to four major questions: What were Saudis' Twitter-based responses and opinions toward the entertaining activities and events organized by the GEA? How could computational linguistic methods be used for offering a reliable conceptual analysis of the Twitter-based responses and opinions voiced regarding such activities? What are the major themes (macropropositions) that framed a polarity in the Saudi public's opinion about the recent entertaining policies? What are the macro-strategies that Saudis implement to polarize discourse on Twitter?

The significance of this study resides in the efficacy of opinion-mining tools in building better information-access systems that would give a clear vision for officials and policymakers about the progress of their policies. In other words, exploring people's reactions, attitudes, and responses toward new policies and plans, especially through social media, might offer insights into their (in)efficacy. Accordingly, they might be either improved or discarded. A part of the significance of this study is that it approaches entertainment policies in the Saudi context, being one of the newsworthy topics that has raised hectic debates locally and regionally in the last four years. Though diverse safety measures have been taken by Saudi officials to ban any crowds to stop the spread of COVID-19, some entertainment events were excluded from such a ban. As mentioned earlier, the exploration of popular responses and sentiments against such policies did not receive the due research interest in Arabic, thereby forming a research gap that the current study sets out to fill.

The rest of the present paper is outlined as follows: Section 2 offers a discussion of previous studies on opinion mining, especially in the Saudi context, and the theoretical framework adopted in this paper. Section 3 offers a detailed description of the corpus design, data collection methods, and data analysis procedure. Section 4 conducts a computational linguistic analysis of the collected tweets based primarily on vector space classification (VSC). Section 5 states the research findings and offers avenues for further research.

2. Literature Review

Nowadays, social networks are gaining much power as they become an integral part of our daily routine. With millions of users having access to diverse social networking applications, most of the official authorities and organizations worldwide established their official accounts to interact with and respond to users' inquiries and opinions (Gunter et al., 2016). Accordingly, policies and decisions are revisited, and moves are redirected. To address and mine billions of messages expressing diverse reactions and responses regarding myriads of political, social, economic, and cultural issues, computational tools have been established and implemented to survey them. Such 'opinion mining' is particularly concerned with the analysis of human behavior expressed via the platforms and networks operating using computer modeling.

Given the focus of the current research, we affirm that a good number of studies have been conducted to analyze Twitter-based sentiments in Arabic (Abdulla et al., 2013; Abu Kwaik et al., 2020; Alrahabi, 2014; Altawaier & Tiun, 2016; Emam & Alzahrani, 2017; Abuelenin et al., 2018). Most of these studies sought to test the classification accuracy of different algorithms (e.g., Support Vector Machine, Decision Tree, K-Nearest Neighbor, and Naive Bayes) and the polarity of opinions regarding a plethora of issues. In the Saudi context, a few studies have analyzed Saudi issues and concerns utilizing opinion mining analytic procedures (Al-Rubaiee et al., 2015; Al-Thubaity et al., 2018; Aldayel & Azmi, 2016; Alliheibi et al., 2021; Alwakid et al., 2017; Azmi & Alzanin, 2014; El-Beltagy & Ali, 2013; Muhajab & Ghazinour, 2019).

El-Beltagy and Ali (2013) highlighted some of the major problems lowering the pace of research in the area of opinion mining in Arabic. The findings revealed that the most significant issues include a lack of colloquial Arabic parsers and sentiment lexicons, difficulty identifying opinion holders whose names are mistaken for sentiments (e.g., نبيل , عادل, and حكيم), and difficulty handling fixed expressions. In the same regard, Aldayel and Azmi (2016) were much concerned with finding a solution to some of the problems associated with the opinion mining of tweets in Saudi Arabic. They found a piece of empirical evidence that combining lexicon classifiers with machine learning classifiers renders a more accurate classification of opinions (positive, neutral, and negative). Relatedly, and towards finding a solution to dialectical issues in the process of sentiment analysis of Arabic opinions over Twitter using either machine learning or lexicon-based techniques, Aldayel and Azmi (2016) offered a hybrid approach incorporating both machine learning techniques and semantic orientation. The data obtained from the lexical classifier is used as training data for the VM machine learning classifier. Findings affirmed that both the F-measure and accuracy of the lexical classifier were significantly improved.

More specifically, Azmi and Alzanin (2014) aimed at the classifications of readers' comments in colloquial Arabic on different issues addressed in the online editions of Saudi newspapers, namely Alriyadh and Aljazirah. They proposed Aara' an opinion mining system using a naive Bayes classifier for extracting the public opinion polarity, focusing on comments introduced in the Arabic Nejdi dialect. Findings showed that the Aara system managed to classify such opinions into four categories (strongly positive, positive, negative, and strongly negative) with an accuracy rate of 82%.

Al-Rubaiee et al. (2015) explored the opinions of Saudi investors as displayed on Twitter regarding the Saudi market index in order to help non-Saudi investors reach a better evaluation of the efficacy of investing in the Saudi market. Using a machine learning approach, they targeted both standard and Arabian Gulf Arabic. For the classification of opinion polarity, three key classifiers were used: Naive-Bayes, SVM, and KNN. The results confirmed the importance of text preprocessing in opinion classification. The SVM and KNN algorithms were much more accurate.

Similarly, exploring the issue of unemployment in Saudi Arabia, Alwakid et al. (2017) addressed some of the challenges of doing opinion-mining research on Arabic social networks with special regard to Saudi dialects, especially the Hejazi and the Nejdi dialects. Saudi dialects, in general, do not strictly follow the syntactic and grammatical rules of Modern Standard Arabic. Furthermore, there are no available, reliable tools for extracting Arabic sentiments. Using linguistic pre-processing of raw texts and machine learning tools, findings showed that the implementation of diverse yet compatible computational techniques (including Naive Bayes and Support Vector Machines) can yield satisfactory sentiment classification accuracy.

More technically, Al-Thubaity et al. (2018) sought to test the efficacy of SauDiSenti, a sentiment lexicon, in sorting out sentiments in Saudi Arabic tweets. Towards this objective, precision, recall, and F measures were used. Furthermore, the results obtained from SauDiSenti were compared to those obtained from AraSenTi, a large Arabic sentiment dictionary. Findings affirmed that, as far as positive, negative, and neutral tweets are regarded, SauDiSenti outperformed AraSenTi.

Not surprisingly, different studies sought to inform decision-makers about the extent to which Saudis are satisfied with new services and policies. For instance, Alsulami and Mehmood (2018) aimed at categorizing Twitter users' attitudes, as well as their expectations, toward the new university system offered by the Ministry of Education in 2018. For this purpose, the SAP HANA platform was used for preparing, transforming, merging, and visualizing data. Findings showed that users had more negative attitudes. Furthermore, regarding users' expectations about the new university system, three types of changes were figured out: administration-related, student-related, and faculty-related. Similarly, Muhajab and Ghazinour (2019) conducted a sentiment analysis of tweets in Saudi dialects discussing women driving in Saudi Arabia. Also, they were concerned with They relied on Weka 3.8 and MonkeyLearn software for data extraction and classification. Findings showed a slight discrepancy in the accuracy of the used software due to the peculiar nature of the lexicon of Saudi Arabic.

More recently, and amid the COVID-19 crisis, Aljameel et al. (2021) aimed at supporting decision-making in the medical sector by investigating differences among Saudi regions regarding their opinions about the precautionary procedures taken since the outbreak of the COVID-19 virus. Employing a Support Vector Machine (SVM), a machine-learning predictive model, the study explored 242,525 tweets. People's awareness was measured based on positive and negative tweets. Findings showed that the southern region of Saudi Arabia had the highest awareness level (65%), while the middle region had the lowest (28%). Relatedly, Aljabri et al. (2021) explored the acceptance rate of new distance learning policies in Saudi Arabia amid the COVID-19 crisis, as exemplified by Saudis' attitudes expressed on Twitter. Different machine learning classifiers were used and then compared. Findings showed that the Logistic Regression classifier was the most accurate. At the general education level, users had positive opinions regarding distance learning. Whereas, at the university level, they had negative opinions.

In a similar context, Alliheibi et al. (2021) explored Saudi responses to the procedure and efficacy of the COVID-19 vaccination policy as documented on Twitter. The corpus included 37.467 tweets collected using four keywords were identified: لقاح كورونا (Corona virus vaccine), تطعيم كورونا (Corona virus vaccine), التحصين (vaccination), and الحصول علي اللقاح (vaccine uptake). Using the computational lexical-semantic analysis that is based on vector space classification, findings showed that most of the Saudi opinions regarding the COVID-19 vaccine take-up calls are markedly negative (54%), reflected in the frequency of negative lexemes such as مؤامرة (conspiracy), جلطات (clots), and عقم (infertility).

Taking into consideration the major theme of the current research (i.e., GEA activities), Alkhaldi et al. (2020) analyzed Saudis' Twitter-based opinions regarding the GEA activities employing three machine learning algorithms (Multilayer Perceptron, SVM, and Random Forest) and Recurrent Neural Network, a deep learning algorithm. Findings showed that most Twitter users support GEA activities, and the majority of them are females. Also, the Random Forest algorithm has been proven to be more accurate and precise in the process of opinion classification.

Though there is a notable burst of research activity in the area of opinion mining, research in Arabic opinion mining has not received due attention, despite the fact that the number of microblog users is increasing rapidly. Also, the most recent steps taken by the Saudi authorities to enhance the quality of life in Saudi Arabia, in parallel with the implementation of Saudi Vision 2030, have not been addressed. Equally important, the available literature is confined to quantitative techniques, while the qualitative dimension—in the sense of interpreting the sociocognitive dimension underlying the resultant classification—is unfortunately absent. Therefore, the current study seeks to help with the exploration of the feasibility of implementing hybrid computational linguistic techniques to identify Saudis' attitudes and responses regarding the GEA activities taking place in the kingdom. The qualitative analysis will be augmented using Van Dijk's notion of the 'ideological square' explained in the following section (Van Dijk, 1993; 2008; 2009).

3. Theoretical Framework

An opinion is claimed to comprise a target with particular attributes and sentiments toward that target (Hu & Liu, 2004). For instance, a sentence such as 'The TV sound is clear, but the resolution is bad' features one target (a product) with two attributes (sound and resolution) expressing positive and negative sentiments, respectively. Nasukawa and Yi (2003) and Jindal and Liu (2006) classify opinions into regular and comparative opinions. A regular opinion expresses a sentiment regarding a particular entity (i.e., a service, a topic, a product, an issue, an event, a person, or an organization) or one or more of its attributes. A regular opinion could be direct (e.g., *the durability of the spare parts is great*) or indirect (e.g., *after using this tool for more than a month, I would say it never helped me*). Conversely, a comparative opinion lists one or more of the mutual attributes of two or more entities (e.g., *this dish is the best in your restaurant*).

Due to the popularity of opinion-rich resources, scholars across various disciplines such as politics (Chen et al., 2010; Yano & Smith, 2010), marketing (Liu et al., 2007), public opinion (O'Connor et al., 2010), and business (Hegde et al., 2015; Park & Javed, 2020) have worked out different approaches and methods for examining people's opinions and sentiments automatically to help in making a decision and revisiting current policies. These approaches and methods are commonly grouped under the generic term of 'social media analytics' which focuses on the way social media contents are used to offer a qualitative analysis of particular issues across various domains (Figure 2). It includes diverse enterprises and techniques such as sentiment analysis, social media classification and filtering, and engagement tracking (Wittwer et al., 2017).

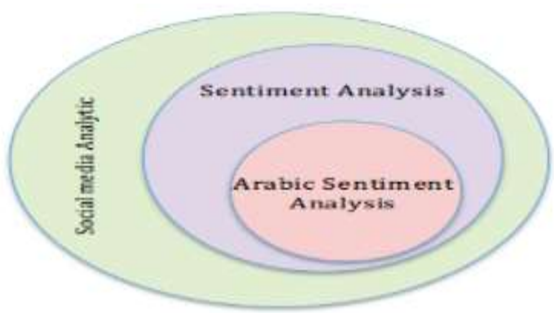

Figure 2. Social media analytic techniques (Alsulami & Mehmood, 2018)

Sentiment analysis evolved as an independent area in computational linguistics, information retrieval, and natural language processing that comprises different tasks such as opinion mining, sentiment analysis, and sentiment orientation (Gupta & Bhathal, 2018; Prabowo & Thelwall, 2009). Indeed, most of the available studies tend to use the terms 'opinion mining' and 'sentiment analysis' interchangeably (Li et al., 2018). Yet, some scholars affirm that 'sentiment analysis' is the computational-based task of classifying opinions and reviews according to their polarity, viz., polarity detection (positive, negative, and neutral) (Pang & Lee, 2008).

Indeed, the history of opinion mining could be traced back to the early 2000s, with a surge of publications targeting the analysis of opinions and sentiments. Similar earlier work sought to interpret metaphors, subjectivity, and viewpoints (Hearst, 1992; Wiebe, 1990). In theory, 'opinion mining', firstly proposed by Dave et al. (2003), refers to the automatic process of detecting, extracting, and classifying opinions, sentiments, responses, appraisals, and judgments posted as texts on the social media regarding a particular issue, law, product, individual, event, service, or policy (Cambria, 2016; Dashtipour et al., 2016; Liu, 2012). Such opinionated data could be found on social media networks such as Twitter, Facebook, Weibo, and forum discussions. In general, users' opinions could be classified as positive, negative, or neutral.

Aside from conducting surveys and opinion polls, it has been empirically affirmed that opinion mining offers a complete picture of people's opinions that could aid in various significant activities such as political voting, stock market prediction, monitoring public opinion, the introduction of new services, and the like. Accordingly, it would facilitate data analysis, planning, and decision-making on the part of individuals as well as organizations across different domains, including politics, public services, healthcare, and consumer services (Li et al., 2018; Liu, 2012).

In general, techniques of opinion mining are claimed to concentrate on three aspects: (1) intentions (e.g., interested or disinterested); (2) polarity of opinions (e.g., positive, neutral, or negative); and (3) personal feelings. In this regard, Dave et al. (2003) propose that an ideal tool in the process of opinion mining would help to reach a set of search results concerning a given item through the generation of a list of features and attributes characteristic of a product as well as opinions about each of such features and attributes (e.g., poor, mixed, good). Though the opinion lexicon is extremely effective at categorizing opinions based on words (e.g., good, bad, wonderful, interesting, boring, etc.) and phrases (e.g., it's just like hell, cost an arm and a leg, etc.), opinionated data could be analyzed at the sentence, document, entity, and aspect levels. At the sentence level, a sentence is said to be either objective or subjective, both of which express positive, negative, or neural opinions (Wiebe et al., 1999). At the document level, a document is assumed to express an overall positive or negative opinion regarding a particular entity (Turney, 2002). At the entity or aspect level, the analysis targets the opinion itself rather than how it is linguistically formed in terms of words, phrases, clauses, and sentences.

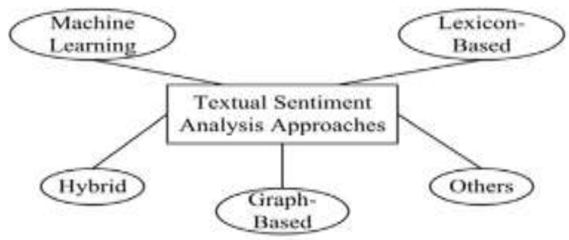

Figure 3. Sentiment analysis approaches for social media text

In textual sentiment analysis, the two essential approaches employed are the machine learning approach and the lexical-based approach (O'Connor et al., 2010; Singh & Husain, 2014). On the one hand, the mechanism of the machine learning approach relies on extracting features from an input text using various mechanisms such as Bag-of-Words, POS taggers, and n-grams. The extracted features are then converted into vector machines to represent useful information. The machine learning approach is mainly used in the task of subjectivity classification through labeling data by sentiment using tools such as Naïve Bayes classifier, SVM, Random Forest, and Logistic Regression (Altawaier & Tiun, 2016). However, the time-consuming task of training samples for annotation is a significant disadvantage of using supervised machine learning techniques. To automate such an annotation process, it is conducted via a bootstrapping technique that employs a replacement technique for stimulating samples out of a dataset (Davidson & Hinkley, 1997).

On the other hand, based on words annotated by polarity, the lexicon-based approach sets out to predict the general sentiment orientation of textual messages based on a pre-created sentiment lexicon of positive, negative, and neutral lexemes constructed either manually or automatically. One key advantage of lexicon-based approaches is that they do not require training data (Taboada et al., 2011). Still, the two techniques could be used simultaneously (Khan et al., 2014).

The process of opinion mining has been empirically proved to comprise a six-step procedure: (1) text pre-processing; (2) feature selection; (3) training set selection; (4) knowledge discovery and presentation; (5) classifier design; and (6) development and evaluation (Subhashini et al., 2021). Firstly, in the step of text pre-processing, all irrelevant information is removed to enhance the classifier's performance, either through tokenization, stemming, or stop-word removal. Secondly, in the process of feature selection, the key features of opinionated data are identified for more effective classification. As shown in Figure 4, such features might be structural, linguistic, or relational. Thirdly, training sets in the form of labeled data are selected to design the classifier. Fourthly, hidden knowledge is identified and presented through different methods, including n-grams, pattern mining, part-of-speech tagging, ontological tagging, and fuzzy logic. Fifthly, a sentiment (supervised or unsupervised) classifier is designed for clustering opinions based on the selected features and the discovered knowledge. Finally, the sentiment classifier is evaluated with reference to benchmarks.

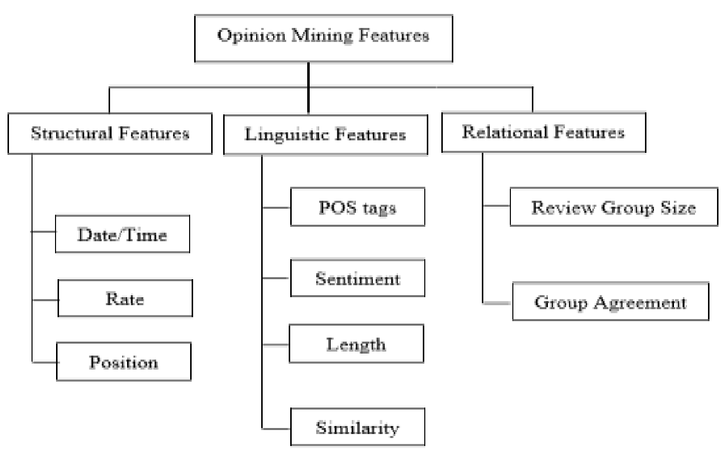

Figure 4. Major features explored in opinion mining (Subhashini et al., 2021)

Still, as affirmed through related literature, opinion mining is claimed to have some caveats**,** especially when addressing sentiments in Arabic (Al-Braheem & Al-Khalifa, 2012; Abuelenin et al., 2018). In general, the tone of the sentiment is not clear, and therefore subjective and objective responses cannot be easily identified. Also, the lexicon used in responses is not always polar, such as 'love' and 'hate', i.e., some phrases could be classified as mid-polar, e.g., 'not so bad'. More interestingly, in the absence of context, sarcastic and ironic comments could be erroneously classified into positive or negative responses based on their surface structure. Given Arabic as the target language to be analyzed in the current paper, sentiments are expressed in dialectical or vernacular Arabic, i.e., Modern Standard Arabic is not the norm. Therefore, phonological, morphological, lexical, orthographic, and syntactic variations abound. Furthermore, loan words and neologisms are sometimes used for fulfilling diverse functions. Also, Twitter users tend to tweet randomly, not caring for grammar and spelling thereby, generating noise in the text (Khasawneh et al., 2013).

Notably, the role of sentiment analysis stops after testing the efficacy of the classifiers and the categories of opinions and attitudes in terms of frequency. In other words, the reasons beyond the statistical differences among positive, negative, and neutral opinions remain unexplored. Therefore, taking into consideration that Twitter forms a particular type of discourse, offering a sociocognitive reading into the output categories or clusters would give many insights to decision-makers as the statistics would be qualitatively justified. In this paper, we rely on Van Dijk's *'ideological square model'* (Van Dijk, 1993) for discussing the reasons underlying the expected divide in opinions regarding the vision of the Saudi General Entertainment Authority. In its broadest sense, Van Dijk's ideological square model, as a key analytic device in his sociocognitive approach (Van Dijk, 2008, 2009), seeks to explain the ideologies underlying social attitudes as exhibited in text and talk by highlighting the nexus between society, social cognition, and discourse. Within this model, different social groups' the discursive strategies (referred to *ingroup* and *outgroup*) are best understood in light of four maxims: INGROUP (emphasizing/expressing good practices; deemphasizing/suppressing bad practices) versus OUTGROUP (emphasizing/expressing bad practices; deemphasizing/suppressing good practices). The gamut of these maxims represents a polarized vision on different social and political issues, mostly motivated by the social group's habitus, values, norms, and conventions, through which other social groups are framed.

4. Methodology

The rationale beyond the choice of Twitter as the context of this study is that Twitter is widely regarded as the most popular social networking site used by the Saudis to express their viewpoints and reflect on current affairs. The target data set is confined to only the tweets written in Arabic, namely Modern Standard Arabic and Saudi Arabic, as the most widely used varieties used by Saudi users of Twitter. Though NLP applications have been largely criticized for not working properly with colloquial varieties of Arabic, the introduction of different social networking sites has caused unprecedented changes in the field.

In view of the research questions raised in the introduction, we designed a corpus of 9378 Twitter posts (tweets), covering the period between 2020 and 2022, collected using Trackmyhashtag, which is one of the social media management platforms allowing hashtag tracking. In this regard, the related literature affirmed the efficacy of hashtags as predictors of the tweet-based sentiments (Simeon & Hilderman, 2015). This hashtag tracking tool serves to search for relevant tweets based on keywords forming a hashtag by specifying language and location. Given the context of this study, the final data set is generated based on five hashtags, as indicated in Table 1.

Table 1. The main hashtags used as entry points for aggregating the final dataset

Hashtags	Transliteration	English translation
هيئة الترفيه	hayyat altarfih	Entertainment Authority
موسم الرياض	mawsim alriyad	Riyadh Season
موسم جدة	mawsim jida	Jeddah Season
موسم الشرقية	mawsim alsharqia	Sharqiah Season
موسم الدرعية	mawsim aldireia	Diriyah Season

In the area of opinion mining, diverse methods, either rule-based or machine learning techniques, are implemented. On the one hand, towards the identification of language users' sentiments and determining their messages' polarity, rule-based methodologies rely on the manual definition of a group of rules in light of which groups of words and phrases signifying particular sentiments are anchored. Polarity is typically identified using two clusters of words. While the first cluster incorporates words with positive connotations such as '*great'*, '*fantastic'*, '*legendary'*, '*good'*, '*robust'*, '*marvelous'*, etc., the second cluster incorporates words with negative connotations such as '*unreliable'*, '*bad'*, '*failing'*, '*deplorable'*, '*repulsive'*, etc. Once these words are identified, they are calculated since their total number would help with determining the polarity of the sentiments. Usually three clusters are generated: positive (+) if words with positive connotations are the most frequent, negative (-) if words with negative connotations are the most frequent, and neutral (±) if the total of both positive and negative words is roughly equal.

On the other hand, machine learning methodologies set out to cluster available data, and they are not driven by any prior assumptions about the datasets. That is, they mainly perceive opinion mining as a data clustering issue where algorithms, rather than hand-written rules, are employed. The process of text clustering, which is an efficient computational system for performing data mining tasks (e.g., summarization, feature extraction, and annotation), is done by means of machine learning algorithms that are manipulated to analyze the informational content of natural language texts so as to group them automatically. The diverse clustering systems can be collectively explained under the umbrella term 'vector space clustering' which is mainly geared toward measuring the semantic similarity of the natural texts (or documents) to be clustered. In this study, vector space clustering is implemented in view of the lexical and semantic features of the linguistic verbiage or content of the dataset covering Saudis' perceptions of the Saudi GEA activities. The procedure of vector space clustering goes through the following steps: (a) text preprocessing; (b) removing function words; (c) stemming; and (d) data representation.

A. Text preprocessing

Towards the transformation of an unstructured textual dataset incorporating the collected tweets into a structured format, we tokenized it, i.e., we deciphered it into a list of tokens or chunks such as words, subwords, phrases, symbols, etc., with the main objective of discarding personal data, non-alphabetic characters, and punctuation marks as well as having analyzable discrete elements. The result of such a step is the establishment of a bag of words (BoW) that does not consider word order, word structure, or context. Words' presence is marked as a boolean value: (1) for present and (0) for absent.

B. Removing function words

Since the process of text clustering typically relies on keyword indexing (i.e., spotting index terms or content words) of the textual dataset, any irrelevant terms are discarded. Though occurring frequently in texts, stop (or function) words such as pronouns, articles, modals, and auxiliaries are counted as irrelevant terms, and therefore they are filtered out since they do not have any semantic weigh that is essential to the processing of the target texts. This step is perceived as complementary to tokenization in preparing the textual dataset to be analyzed by a machine learning model with the main focus on high-level information, thereby reducing the training time for text mining.

C. Stemming

As soon as the index terms are listed and the stop words are filtered out, the stemming process can begin. By stemming, we refer to the use of language-dependent algorithms to reduce words to their stems or roots by removing inflectional or derivational morphemes to conflate semantically related words and word variants. The rationale beyond stemming is that derivational and inflectional morphemes are not semantically rich. Accordingly, their removal is essential in almost all text clustering software. In so doing, all and only words that are semantically related and share the same stem are grouped. Equally important, conflation enhances the performance of clustering systems as data size and complexity are markedly reduced. However, given the inflected nature of Arabic, we use a light-based stemmer rather than a root-based stemmer to remove only prefixes and suffixes while keeping other necessary morphological information (i.e., infixes) that helps to reach a better understanding of the dataset profile. Therefore, we use Light10 stemmer, as it has proven highly effective in Arabic data retrieval (Larkey et al., 2007).

D. Data representation

Having stemmed and conflated the collected tweets, the next step is concerned with data representation. Toward this objective, a row vector matrix is built and structured as (Xij), representing data dimensions. That is, where (X) represents the vector, (i) represents the tweets (rows), and (j) represent the lexical types (columns). It offers a quantitative overview of the dataset in terms of word frequencies, types, and tokens. Yet due to the fact that tweets vary in terms of their length, they must be penalized or normalized (i.e., to reduce texts' randomness) to guarantee that every single tweet contributes equally and significantly to the distance in vector space clustering, thereby ensuring that all distances are equal and the clustering process is reliable. The rationale is that long tweets containing more words or more contents have a better chance of matching the target queries. The non-normalization of tweets might cause longer tweets to dominate the distance calculation. Therefore, to render all tweets (on the row vectors of the matrix) to be of average length to improve the similarity function, we used cosine normalization based on the following formula:

$$Freq = Freq\ (F_{ij})\frac{\mu}{length(i)}$$

As all tweets are normalized, there remains one challenge hurting the reliability of the automatic classification systems: high dimensionality. Referring to the data as high-dimensional means that the number of features observed is greater than (or even equal to) the number of observations. If the number of such features is markedly high, the process of classifying documents in view of their semantic properties becomes difficult since there are not enough observations to train the vector space clustering model. Therefore, we used a statistical measure known as TF-IDF (frequency-inverse document frequency) to reduce the number of features being observed by considering only distinctive features. As a result, the final dataset was reduced to the highest 200 variables, incorporating highly frequent relevant words (Figure 4). Tweets with similar relevant words have similar vectors.

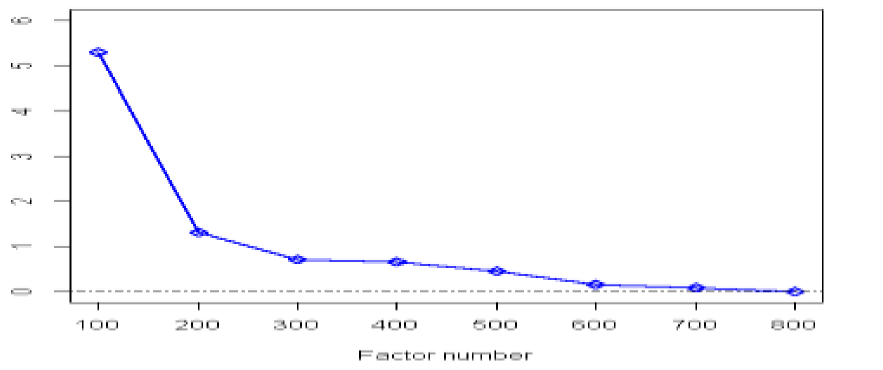

Figure 4. The highest 200 variables in the dataset

Finally, we clustered the tweets hierarchically by means of Euclidean distance methods by calculating the (dis)similarity between pairs of observations.

E. Analysis & Discussion

To analyze the final data matrix hierarchically by means of Euclidean distance as well as squared Euclidean distance measures, we initially used four clustering methods: (1) single linkage, (2) complete linkage, (3) average linkage, and (4) Ward's linkage (also known as *'sum of squares'*). The rationale for implementing the four methods is that there is a consensus that one clustering algorithm is not a panacea. After the implementation of the four methods, the output clusters are eventually compared for (dis)agreement on the data structure, resulting in more valid and reliable clusters. Findings affirmed that there is no statistically significant difference between Euclidean and squared Euclidean measures as they produced similar proximity metrics incorporating the main clusters, though the detailed structure was different across the four methods. Further to this, the output clusters were much clearer through Ward's linkage clustering method, which worked best with Euclidean distance measures.

Though the four methods implemented in the process of textual clustering in this study are generally classified as agglomerative hierarchical clustering methods, their output clustering structures varied in terms of clustering clarity and efficacy. Firstly, the single linkage clustering method has been proven to be ineffective in identifying meaningful clusters due to chaining. That is, by chaining the tweets together, the resulting structure is noisy and obscured, making it difficult to define classes that could aid in data subdivision. Accordingly, subjectivity issues might hurt the reliability of the produced clusters. Contrarily, the complete linkage clustering method offered a clearer subdivision of the dataset; it broke large clusters, thereby producing a large number of small-membership clusters.

Thirdly, though the average linkage clustering method is proposed to overcome the problems associated with single and complete linkage clustering methods, it did not manage to offer clear clusters of the target tweets owing to its overwhelming tendency toward left branching. Additionally, the resultant clusters were inconsistent, viz., very small clusters versus very large clusters. Finally, as mentioned earlier, the output clusters were much clearer and better-partitioned through Ward's linkage clustering method. That is, the responses of Saudis regarding the activities of the GEA were clearly subdivided into two well-identified clusters, as illustrated in Figure 5.

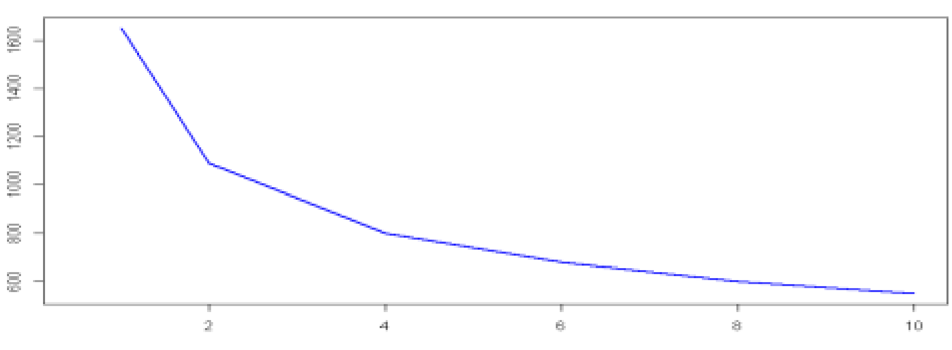

Figure 5. Identification of the number of clusters

Despite the differences in the clustering structures produced by the implemented clustering methods, they all agreed on splitting the dataset into two clusters or groups: Group A and Group B.

F. Centroid comparison

Once the dataset of tweets is divided into two clusters, marked 'Group A' and 'Group B', representing different responses and attitudes regarding the GEA activities, we conducted centroid analysis to identify the difference between the two vectors representing both groups over the 200 previously identified variables in numerical values. Such centroid analysis identifies the arithmetic mean of all the data points belonging to each cluster. It also determines the most and least frequent variables that characterize each cluster. Table 2 below shows the difference between the centroid frequencies of Group A and Group for the most frequent variables.

Table 2. Centroid comparison of Group A & Group B

Variable index	**Centroid 1 frequency Group A**	**Centroid 2 frequency Group B**	**Difference**
18	1.66421E009	6.84766E009	5.18346E009
7	1.8777E008	1.45939E009	1.27162E009
4	1.21382E009	0.000	1.21382E009
21	1.46071E008	1.23528E009	1.08921E009
2	1.13932E009	2.19306E009	1.05374E009
6	1.76854E008	1.08254E009	9.05687E008
9	2.81168E008	1.04827E009	7.67101E008
14	1.01934E009	2.67894E008	7.5145E008
19	0.000	7.0597E008	7.0597E008
17	1.22763E008	7.89109E008	6.66345E008
40	2.21264E008	8.84367E008	6.63103E008
8	2.24012E009	1.58934E009	6.50773E008
10	6.596E007	6.21396E008	5.55436E008
0	7.34327E008	1.82779E008	5.51548E008
24	5.13532E008	0.000	5.13532E008
31	5.05069E008	1.81161E007	4.86953E008
22	7.09206E008	2.31899E008	4.77307E008
26	7.86125E008	1.22338E009	4.37256E008
80	1.74257E008	5.71459E008	3.97202E008
43	2.0996E007	4.05783E008	3.84787E008
88	7.19377E008	3.42165E008	3.77212E008
112	7.65595E008	3.91658E008	3.73936E008
42	1.73487E009	1.40337E009	3.31492E008
52	7.06457E007	3.75202E008	3.04556E008
55	1.01694E008	3.94238E008	2.92544E008
11	9.34589E008	6.44466E008	2.90123E008

Results of the centroid comparison revealed that there is a statistically significant difference between the members of Group A and Group B. A close reading into the distinctive variables in each group showed that Group A members, on the one hand, approve of the activities, events, and festivals organized and run by the GEA. Table 3 below displays some of the keywords indicating consensus and agreement.

Table 3. Distinctive lexical features of group A

Group A		
Arabic words	**Transliteration**	**English translation**
حلوة	Hulwa	sweety
هايل	Hayil	wonderful
ناجح	Najih	successful
روعة	Raweah	terrific
موطبيعي	mo tabiei	Abnormal
معقول	Maequl	Fair
خيالي	Khayali	legendary
مبهجة	Mubhija	cheerful
رائع	Rayie	terrific
كفو	Kafu	perfect
حلو	Hulw	Good
مرَّة	Marrah	extremely
عظيم	Azim	Great
ابداع	Ibdae	creativity
أحلى	Ahlaa	much better
اتبسطنا	Itbasatna	we are pleased

On the other hand, Group B members disapproved of such activities and events. Some of the distinctive lexical features expressing disapproval and disagreement of the GEA activities are shown in Table 4.

Table 4. Distinctive lexical features of Group B

Group B		
Arabic words	**Transliteration**	**English Translation**
لا	La	No
فاشل	Fashil	failure
قاطعوا	Qataeuu	boycott
أوقفوا	Awqifuu	Stop
مانبغى	ma nabghaa	we reject
ضد	Did	against
أبد	Abad	never

According to the results of the centroid analysis, the total number of tweets in each cluster varied considerably. While the tweets of Group A represented 62% of the dataset, Group B's tweets represented only 38%. Therefore, Given the limitations of the data, it could be concluded that Saudis have positive attitudes towards all the vibes and events organized and sponsored by the Saudi GEA. However, such a discrepancy in numerical values represents only a rough idea of the discrepancy in Saudis' sentiments and opinions on GEA-organized events. In other words, the findings of the clustering process remain insufficient due to inherent limitations in machine learning algorithms. Therefore, further analysis of the tweets' discourse features, realizing the polarity in the dataset, remains highly significant to augment the statistical analysis and reach a better understanding. Moreover, contextualizing the distinctive lexical features of Group A and Group B offers an interpretative framework for Saudi users' opinions and sentiments mediated through Twitter.

To computationally identify the salient discourse features of the tweets, keyness and lexical associations should be established. Toward this objective, five keywords—representing the main hashtags under investigation—were selected and tabulated below.

Table 5. Frequency of key words

No.	Keyword	English Equivalent	Frequency
1	هيئة الترفيه	GEA	4826
2	موسم الرياض	RIYADH SEASON	1791
3	موسم جدة	JEDDAH SEASON	1137
4	موسم الشرقية	SHARQIAH SEASON	719
5	موسم الدرعية	DIRIYAH SEASON	558

For this reason, we used the concordancing tool known as Key Word in Context (KWIC) to generate concordances through which the context underlying the target keywords could be revealed. The tweets pertaining to each group are analyzed separately to be easily compared. On the one hand, the tweets clustered in Group A were found to be associated with themes with positive associations such as 'new kingdom', 'modern country', 'fighting terrorism', 'legal supremacy', 'economic reform', 'more jobs', 'great Saudi Arabia', etc. For more keywords' associations, see Table 6.

Table 6. Associations of keywords in Group A

Collocates	Translation	Frequency
سيادة القانون	rule of law	1453
المعتدل	moderate	1267
حقوق المواطن	citizens' rights	988
محاربة التطرف	fighting extremism	1124
نبذ العنف	renouncing violence	976
رؤية السعودية 2030	Saudi vision 2030	2365
مجتمع منفتح	open society	2234
أكثر انفتاحا	more open	1876
عصري	modern	1493
طموح	ambitious	657
تحول كبير	great transformation	765
السعودية العظيمة	Great Saudi Arabia	884
وظائف أكثر	more jobs	659
إصلاح اقتصادي	economic reform	458
الإسلام المعتدل	moderate Islam	523
توظيف الشباب السعودي	employing Saudi youth	466
فرص عمل	job opportunities	398
استثمارات	investments	411
قرارات إصلاحية	corrective decisions	396
الانفتاح	openness	485

On the other hand, the five key words GEA, Riyadh season and jeddah season, SHARQIAH SEASON, and DIRIYAH SEASON were found to be associated with key negative concepts including corruption, Muslim identity, and traditions. Table 7 below displays more words and phrases with negative associations.

Table 7. Associations of keywords in Group B

Collocates	Translation	Frequency
فساد	corruption	1287
انحلال	decadence	1165
سفور	bareheadedness	834
اختلاط الرجال والنساء	mixing of men and women	766
تقاليد	traditions	1136
الهوية	identity	1976
بلاد التوحيد	country of monotheism	886
المجون	obscenity	455
التحرش	harassment	399
الطابع الديني للمملكة	the religious nature of the Kingdom	377

As demonstrated in tables 7 and 8, there is a significant polarity in the dataset representing Saudis' opinions on the activities and vibes organized by the GEA. That is, the members of Group A and Group B offered two opposing discourses, representing a clear divide in Saudis' perceptions of such activities. Each group's opinions are motivated by their habitus and preconceived ideological values. In light of the theoretical underpinnings of Van Dijk's ideological square and at a more abstract underlying level of analysis, there is an ideological divide between the two groups instantiated by the manipulation of the collective first-person pronouns 'we/us' (*self/ingroup*) versus the third-person pronouns 'they/them' (*others/outgroup*).

On the one hand, the Group A members' opinionated data showed that their support of these activities is discursively (at the macro level) grounded through four main themes or macropropositions including the Kingdom's Vision 2030, geared toward enacting a more effective economic reform, being transformed into a more open country, securing more job and investment opportunities, and combating extremism and violence. On the other hand, the Group B members' sentiments were ideologically motivated, as demonstrated through the overall organization of their tweets around one major theme: the GEA activities as a threat to and an attack against the Muslim identity of Saudi Arabia. This theme has been rendered textually salient (at the micro level) through the reiteration of labels such as 'corruption', 'decadence', and 'obscenity'.

Based on the representations of both groups, we conclude that Group A ideologically perceives Group B members as hardliners who are against the Kingdom's Vision 2030, which mainly aims for economic reform and securing better life standards. As shown in the following screenshots, advocates of the GEA activities perceive that any criticism leveled against such activities is meant to denigrate the kingdom's touristic reputation and to curb its economic power. They argue that the same entertainment and recreational activities abound in other Arab countries, but opponents of these activities are ideologically motivated to defame the kingdom.

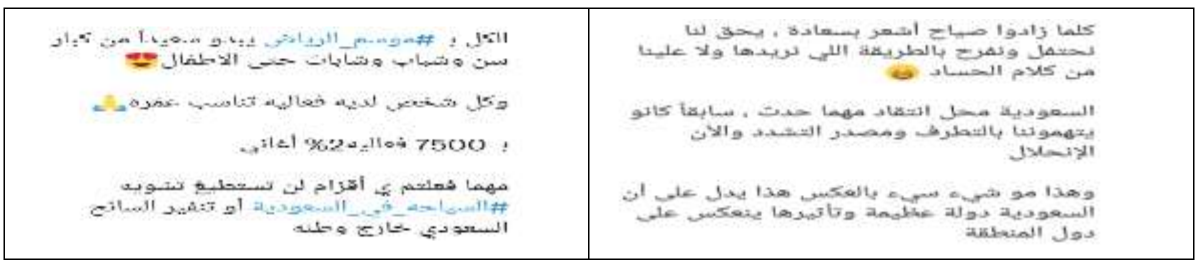

Figure 6. Screenshots of examples of tweets supporting the GEA activities

Conversely, as illustrated in Figure 7, Group B perceives Group A members as secularists who do not abide by the teachings of Islam as well as the Saudi habitus. Accordingly, they believe that the kingdom should take considerable measures to stop any activity that might hurt its privacy, being the cradle of Islam and the incubator of the Two Holy Mosques.

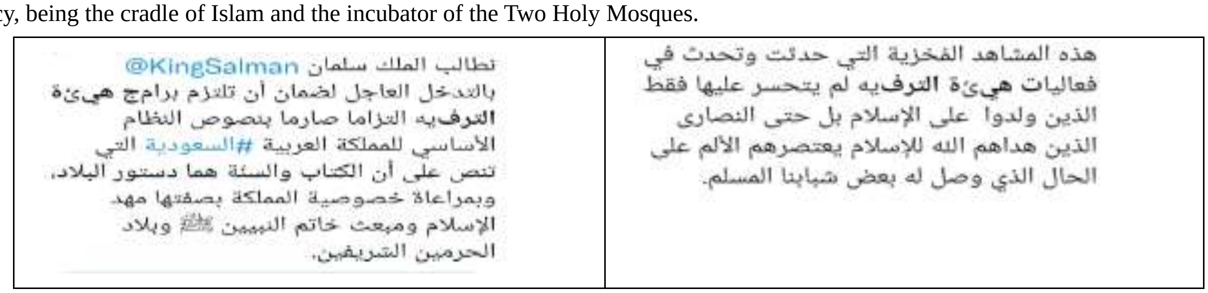

Figure 7. Screenshots of examples of tweets opposing the GEA activities

To sum up, the macrostrategies of positive self-presentation and negative other-presentation employed by both groups are implemented by Saudis as polarization strategies to support their sentiments. Therefore, we can conclude that the computational processing of Twitter discourse managed to offer a clear input to the use of Van Dijk's analytic model known as 'ideological square' to elucidate Twitter behind positive and negative representations of the *self* and *others* since the investigation of discourse features helped to identify particular nuances in the discursive construction of others to confirm group dominance.

5. Conclusion and Future Research

The current study used both computational tools in the form of a machine learning predictive model (Support Vector Machine) and a critical discourse analysis model (Van Dijk's ideological square model) to fulfill three objectives: (1) clustering the Saudis' Twitter-based opinions and sentiments regarding the entertaining and recreational activities run by the Saudi General Entertainment Authority; (2) offering an empirical piece of evidence on how computational linguistic methods could be implemented for offering a reliable conceptual framing of such opinionated data; and (3) outlining the central themes generating a polarity in the Saudi public opinion as well as the macrostrategies through which this polarity is instantiated and actualized. To fulfill these objectives, we built a corpus of 9378 tweets covering the period between 2020 and 2022. The corpus was processed through the different stages of tokenization, conflation, stemming, and representation, leading to the building of the final clusters.

Findings affirmed the efficacy of incorporating the Support Vector Machine model and the ideological square model in clustering and interpreting the target tweets. Based on the output discourse features and thematization of the tweet, two main groups with different ideologically motivated perspectives were highlighted. This ideological polarity was realized through the implementation of two macrostrategies: positive self-presentation and negative other-presentation. It is recommended to conduct future studies to trace such polarity across years to spot any statistically significant differences since the findings could be used by decision-makers to revisit entertainment policies in the Kingdom for better outcomes.

Acknowledgments

The authors extend their appreciation to Prince Sattam bin Abdulaziz University for funding this research work through the project number (PSAU/2022/02/21630).

References

Abdulla, N. A., Ahmed, N. A., Shehab, M. A., & Al-Ayyoub, M. (2013). Arabic sentiment analysis: Lexicon-based and corpus-based. *2013 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT)*, 1-6. https://doi.org/10.1109/aeect.2013.6716448

Abu Kwaik, K., Chatzikyriakidis, S., Dobnik, S., Saad, M., & Johansson, R. (2020). *An Arabic tweets sentiment analysis dataset (ATSAD) using distant supervision and self training.* Paper presented at the Proceedings of the 4th Workshop on Open-Source Arabic Corpora and Processing Tools, with a Shared Task on Offensive Language Detection, Marseille, France.

Abuelenin, S., Elmougy, S., & Naguib, E. (2018). Twitter sentiment analysis for Arabic tweets. In S. K. Hassanien A., Gaber T., Tolba M. (Ed.), *Proceedings of the International Conference on Advanced Intelligent Systems and Informatics 2017*, 467-476. https://doi.org/10.1007/978-3-319-64861-3_44

Al-Braheem, L., & Al-Khalifa, H. S. (2012). Exploring the problems of sentiment analysis in informal Arabic. *Proceedings of the 14th International Conference on Information Integration and Web-based Applications & Services - IIWAS '12*. https://doi.org/10.1145/2428736.2428813

Aldayel, H. K., & Azmi, A. M. (2016). Arabic tweets sentiment analysis–A hybrid scheme. *Journal of Information Science, 42*(6), 782-797. https://doi.org/10.1177/0165551515610513

Aljabri, M., Chrouf, S. M. B., Alzahrani, N. A., Alghamdi, L., Alfehaid, R., Alqarawi, R., . . . Alduhailan, N. (2021). Sentiment analysis of Arabic tweets regarding distance learning in Saudi Arabia during the COVID-19 pandemic. *Sensors (Basel), 21*(16). https://doi.org/10.3390/s21165431

Aljameel, S. S., Alabbad, D. A., Alzahrani, N. A., Alqarni, S. M., Alamoudi, F. A., Babili, L. M., … Alshamrani, F. M. (2021). A sentiment analysis approach to predict an individual's awareness of the precautionary procedures to prevent COVID-19 outbreaks in Saudi Arabia. *International Journal of Environmental Research and Public Health, 18*(1), 1-12. https://doi.org/10.3390/ijerph18010218

Alkhaldi, S., Alzuabi, S., Alqahtani, R., Alshammari, A., Alyousif, F., Alboaneen, D. A., & Almelihi, M. (2020, 19-21 March 2020). Twitter sentiment analysis on activities of Saudi General Entertainment Authority. *2020 3rd International Conference on Computer Applications & Information Security (ICCAIS)*. https://doi.org/10.1109/iccais48893.2020.9096738

Alliheibi, F. M., Omar, A., & Al-Horais, N. (2021). Opinion mining of Saudi responses to COVID-19 vaccines on twitter. *International Journal of Advanced Computer Science and Applications (IJACSA), 12*(6), 72-78. https://doi.org/10.14569/ijacsa.2021.0120610

Alrahabi, M. (2014). Opinion mining from Arabic quotations. *Proceedings of the 6th International Conference on Management of Emergent Digital EcoSystems - MEDES '14*. https://doi.org/10.1145/2668260.2668291

Al-Rubaiee, H., Qiu, R., & Li, D. (2015). Analysis of the relationship between Saudi twitter posts and the Saudi stock market. *Paper presented at the IEEE Seventh International Conference on Intelligent Computing and Information Systems (ICICIS'15), Cairo, Egypt*. https://doi.org/10.1109/intelcis.2015.7397193

Alsulami, M., & Mehmood, R. (2018). Sentiment analysis model for Arabic tweets to detect users' opinions about government services in Saudi Arabia: Ministry of Education as a case study. *Al Yamamah Information and Communication Technology Forum, Riyadh, Saudi Arabia*.

Altawaier, M. M., & Tiun, S. (2016). Comparison of machine learning approaches on Arabic twitter sentiment analysis. *International Journal on Advanced Science, Engineering and Information Technology, 6*(6), 1067-1073. https://doi.org/10.18517/ijaseit.6.6.1456

Al-Thubaity, A., Alqahtani, Q., & Aljandal, A. (2018). Sentiment lexicon for sentiment analysis of Saudi dialect tweets. *Paper presented at the The 4th International Conference on Arabic Computational Linguistics (ACLing 2018)*, Dubai, United Arab Emirates. https://doi.org/10.1016/j.procs.2018.10.494

Alwakid, G., Osman, T., & Hughes-Roberts, T. (2017). Challenges in sentiment analysis for Arabic social networks. *Procedia Computer Science, 117*, 89-100. https://doi.org/10.1016/j.procs.2017.10.097

Azmi, A. M., & Alzanin, S. M. (2014). Aar'- A system for mining the polarity of Saudi public opinion through e-newspaper comments. *Journal of Information Science, 40*, 398-410. https://doi.org/10.1177/0165551514524675

Cambria, E. (2016). Affective computing and sentiment analysis. *IEEE Intelligent System, 31*(2), 102-107. https://doi.org/10.1109/MIS.2016.31

Chen, B., Zhu, L., Kifer, D., & Lee, D. (2010). What is an opinion about? exploring political standpoints using opinion scoring model. Proceedings of AAAI Conference on Artificial Intelligence (AAAI-2010), Atlanta, Georgia, USA, *24*(1), 1007-1012. https://doi.org/10.1609/aaai.v24i1.7717

Dashtipour, K., Poria, S., Hussain, A., Cambria, E., Hawalah, A. Y. A., Gelbukh, A., & Zhou, Q. (2016). Multilingual sentiment analysis: State of the art and independent comparison of techniques. *Cognitive Computation, 8*(4), 757-771. https://doi.org/10.1007/s12559-016-9415-7

Dave, K., Lawrence, S., & Pennock, D. (2003). Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. *Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews, 775152*. https://doi.org/10.1145/775152.775226

Davidson, A. C., & Hinkley, D. V. (1997). *Bootstrap methods and their application*. Cambridge University Press. https://doi.org/10.1017/CBO9780511802843

El-Beltagy, S. R., & Ali, A. (2013). Open issues in the sentiment analysis of Arabic social media: A case study. *2013 9th International Conference on Innovations in Information Technology (IIT) Al Ain, United Arab Emirates*. https://doi.org/10.1109/innovations.2013.6544421

Emam, A., & Alzahrani, M. (2017). Opinion mining techniques and tools: A case study on Arab newspaper. 2017 International Conference on Computational Science and Computational Intelligence (CSCI). https://doi.org/10.1109/csci.2017.49

Gunter, B., Elareshi, M., & Al-Jaber, K. (2016). *Social media in the Arab world: Communication and public opinion in the Gulf states*. Bloomsbury Publishing. https://doi.org/10.5040/9781350988057

Gupta, G., & Bhathal, G. S. (2018). *Sentiment analysis of English tweets using data mining: Data mining, sentiment analysis*. BookRix.

Hearst, M. (1992). Direction-based text interpretation as an information access refinement. In P. Jacobs (Ed.), *Text-based intelligent systems* (pp. 257-274). Lawrence Erlbaum Associates.

Hegde, V., Karthika, P., & Madhu, M. G. (2015). Opinion mining and market analysis. *International Journal of Applied Engineering Research, 10*(10), 25629-25636.

Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. *Proceedings of the 2004 ACM SIGKDD international conference on knowledge discovery and data mining - KDD '04*. https://doi.org/10.1145/1014052.1014073

Jindal, N., & Liu, B. (2006). Mining comparative sentences and relations. *Proceedings of National Conference on Artificial Intelligence (AAAI-2006)*, Boston, Massachusetts, 1331-1336.

Khan, F. H., Bashir, S., & Qamar, U. (2014). TOM: twitter opinion mining framework using hybrid classification scheme. *Decision Support Systems, 57*, 245-257. https://doi.org/10.1016/j.dss.2013.09.004

Khasawneh, R., Wahsheh, H., A., Al-Kabi. M. N., & Aismadi, I. M. (2013). *Sentiment analysis of Arabic social media content: A comparative study. 8th International Conference for Internet Technology and Secured Transactions (ICITST-2013)*. https://doi.org/10.1109/icitst.2013.6750171

Larkey, L. S., Ballesteros, L., & Connell, M. E. (2007). Light stemming for Arabic information retrieval. In A. Soudi, A. V. D. Bosch, & G. Neumann (Eds.), *Arabic comptayion morphology: Knowledge-based and empirical methods* (pp. 221-244). Springer. https://doi.org/10.1007/978-1-4020-6046-5_12

Li, Z., Fan, Y., Jiang, B., Lei, T., & Liu, W. (2018). A survey on sentiment analysis and opinion mining for social multimedia. *Multimedia Tools and Applications, 78*(6), 6939-6967. https://doi.org/10.1007/s11042-018-6445-z

Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies, 5*(1), 1-167. https://doi.org/10.2200/S00416ED1V01Y201204HLT016

Liu, J., Cao, Y., Lin, C., Huang, Y., & Zhou, M. (2007). Low-quality product review detection in opinion summarization. *Proceedings of the*

Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL-2007), Prague, Czech Republic.

Muhajab, H., & Ghazinour, K. (2019). Sentiment analysis of women driving in Saudi Arabia. *Proceedings of ISSAT International Conference on Data Science & Intelligent Systems*, Las Vegas, NV, U.S.A.

Nasukawa, T., & Yi, J. (2003). Sentiment analysis: Capturing favorability using natural language processing. *Proceedings of the international conference on Knowledge capture - K-CAP '03*. https://doi.org/10.1145/945645.945658

O'Connor, B., Balasubramanyan, R., Routledge, B. R., & Smith, N. A. (2010). From tweets to polls: linking text sentiment to public opinion time series. *Proceedings of the International AAAI Conference on Web and Social Media (ICWSM 2010)*, Menlo Park, California, *4*(1), 122-129. https://doi.org/10.1609/icwsm.v4i1.14031

Pak, A., & Paroubek, P. (2010). Twitter as a corpus for sentiment analysis and opinion mining. *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10), Valletta, Malta, 5*(12), 320-322. https://doi.org/10.17148/ijarcce.2016.51274

Pang, B., & Lee, L. (2008). Opinion mining and sentiment snalysis. *Foundations and Trends® in Information Retrieval, 2*(1-2), 1-135. https://doi.org/10.1561/1500000011

Park, Y. E., & Javed, Y. (2020). Insights discovery through hidden sentiment in big data: Evidence from Saudi Arabia's financial sector. *Journal of Asian Finance, Economics and Business, 7*(6), 457-464. https://doi.org/10.13106/jafeb.2020.vol7.no6.457

Prabowo, R., & Thelwall, M. (2009). Sentiment analysis: A combined approach. *Journal of Informetrics, 3*(2), 143-157. https://doi.org/10.1016/j.joi.2009.01.003

Simeon, C., & Hilderman, R. (2015). Evaluating the effectiveness of hashtags as predictors of the sentiment of tweets. *Discovery Science*, 251-265. https://doi.org/10.1007/978-3-319-24282-8_21

Singh, P. K., & Husain, M. S. (2014). Methodological study of opinion mining and sentiment analysis techniques. *International Journal on Soft Computing, 5*(1), 11-21. https://doi.org/10.5121/ijsc.2014.5102

Subhashini, L. D. C. S., Li, Y., Zhang, J., Atukorale, A. S., & Wu, Y. (2021). Mining and classifying customer reviews: a survey. *Artificial Intelligence Review, 54*(8), 6343-6389. https://doi.org/10.1007/s10462-021-09955-5

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. *Computational Linguistics, 37*(2), 267-307. https://doi.org/10.1162/COLI_a_00049

Turney, P. D. (2002). *Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. Proceedings of the 40th Annual Meeting on Association for Computational Linguistics - ACL-2002, Philadelphia, The United States*. https://doi.org/10.3115/1073083.1073153

Van Dijk, T. A. (1993). Principles of critical discourse analysis. *Discourse & Society, 4*(2), 249-283. https://doi.org/10.1177/0957926593004002006

Van Dijk, T. A. (2008). *Discourse and context. A socio-cognitive approach*. Cambridge University Press.

Van Dijk, T. A. (2009). *Society and discourse. How social contexts influence text and talk*. Cambridge University Press. https://doi.org/10.1017/CBO9780511575273

Wiebe, J. (1990). Identifying subjective characters in narrative. *Proceedings of the 13th conference on Computational Linguistics* (COLING-1990). https://doi.org/10.3115/997939.998008

Wiebe, J., Bruce, R. F., & O'Hara, T. P. (1999). Development and use of a gold-standard data set for subjectivity classifications. *Proceedings of the 37th* annual *meeting of the Association for Computational Linguistics on Computational Linguistics* (ACL-1999). https://doi.org/10.3115/1034678.1034721

Wittwer, M., Reinhold, O., Alt, R., Jessen, F., & Stüber, R. (2017). Social media analytics using business intelligence and social media tools – Differences and implications. *Business Information Systems Workshops*, 252-259. https://doi.org/10.1007/978-3-319-52464-1_23

Yano, T., & Smith, N. A. (2010). What's worthy of comment? Content and comment volume in political blogs. *Proceedings of the International AAAI Conference on Web and Social Media, 4*(1), 359-362. https://doi.org/10.1609/icwsm.v4i1.14062