Sentiment Analysis of News Comments: A Comparison of Human and Automated Emotion Detection Methods

تحليل المشاعر في التعليقات على الأخبار:
مقارنة بين الأساليب البشرية والآلية في الكشف عن العواطف

Dr. Ingy Farouk Emara
Associate Professor, Department of AL-Alsun (Languages)
Faculty of Al-Alsun & Mass Communication
Misr International University

د. إنجي فاروق عمارة
أستاذ مساعد بقسم الألسن
كلية الألسن والإعلام، جامعة مصر الدولية

Sentiment Analysis of News Comments: A Comparison of Human and Automated Emotion Detection Methods

This paper conducts a sentiment analysis comparing human and automated sentiment annotation of Facebook comments associated with news articles likely to evoke the emotions of anger, fear, sadness, and happiness. The study finds that both human and automated methods assigned mostly similar sentiment polarities—negative for comments on the articles triggering anger, fear, and sadness, and positive for comments on the article evoking happiness. However, human annotators detected a wider range of emotional words, while the automated tool missed many of them and, at times, provided inaccurate descriptions of emotions. The study also employs Martin and White's (2005) appraisal theory to examine the emotion-related language structures in these comments. It reveals that the affect dimension predominated in discussions of the sadness-related article, the judgment dimension was more prominent in discussions of the anger- and happiness-related articles, and the appreciation dimension featured more in discussions of the fear-related article.

Keywords: human and automated sentiment analysis, emotion detection, appraisal theory

تحليل المشاعر في التعليقات على الأخبار:
مقارنة بين الأساليب البشرية والآلية في الكشف عن العواطف

الملخص:

تُجري هذه الورقة تحليلًا للمشاعر من خلال مقارنة بين التوصيف البشري والآلي للمشاعر الموجودة في تعليقات على مقالات إخبارية منشورة على منصة الفيسبوك والتي يُحتمل أن تثير مشاعر الغضب، والخوف، والحزن، والسعادة. وتجد الدراسة أن كلًا من الطرق البشرية والآلية أسندت بشكل عام اتجاهات مشاعر متشابهة ـــ سلبية للتعليقات على المقالات التي تثير الغضب والخوف والحزن، وإيجابية للتعليقات على المقالة التي تثير السعادة. ومع ذلك، رصد المحللون البشريون نطاقًا أوسع من الكلمات العاطفية، مما أتاح وصفًا أكثر تفصيلًا للمشاعر، بينما تجاهلت الأداة الآلية LIWC-22 العديد من الكلمات المرتبطة بالعواطف، وقدمت أحيانًا أوصافًا غير دقيقة للمشاعر لا تتماشى مع السياق. كما تطبق الدراسة نظرية التقييم بأبعادها الثلاثة لمارتن ووايت (٢٠٠٥) وذلك لاستكشاف الأبنية اللغوية المرتبطة بالعواطف في هذه التعليقات. وقد كشفت الدراسة أن بعد "العاطفة" كان مهيمنًا في النقاشات المتعلقة بالمقالة المرتبطة بالحزن، بينما ظهر بعد "الحكم" في النقاشات حول مقالات الغضب والسعادة، وظهر بعد "التقدير" في النقاشات المتعلقة بمقالة الخوف.

الكلمات المفتاحية: تحليل المشاعر الآلي والبشري، الكشف عن العواطف، نظرية التقييم

Sentiment Analysis of News Comments: A Comparison of Human and Automated Emotion Detection Methods

Introduction

Recent advancements in sentiment analysis and speech emotion recognition technologies have significantly influenced how language is used and interpreted across social media platforms. Sentiment analysis and emotion recognition (or detection) are two areas of natural language processing (NLP) that are increasingly used at the present time to investigate online users' views on topics in various fields such as marketing, health, education, politics, and social issues. The main difference between the two approaches is that sentiment analysis detects whether a text has an overall positive, negative, or neutral tone, while emotion recognition investigates what specific emotions are inherent in a text (Nandwani & Verma, 2021).

The current study aims to perform sentiment analysis and emotion recognition on English-language news comments made by Facebook users of different ethnic backgrounds. The selected comments represent social media users' reactions to topics arousing the four emotions of anger, fear, sadness and happiness that comprise the basic human emotions (Ekman, 1992; Wilson-Mendenhall et al., 2013; Jack et al., 2014; Gu et al., 2019). The users' language is analyzed in terms of the appraisal theory proposed by Martin and White (2005) that outlines the language evaluative functions expressing human emotions.

Social media has lately dominated communication patterns and dissemination of information through different digital media platforms, the most popular of which are Facebook and X (formerly Twitter). Sentiment analysis, also known as opinion mining, of social media texts is a realm that is largely investigated due to the increasing influence of social media on users of all ages and ethnic backgrounds. However, a gap in sentiment analysis research exists in the scarcity of studies comparing human and automated sentiment analysis of comments made by social media users on topics evoking the basic human emotions. Therefore, the present research uses the LIWC-22 (Linguistic Inquiry Word Count) sentiment analysis tool, which performs both sentiment analysis and emotion recognition of text samples using a dictionary of emotion words (Boyd et al., 2022), and then compares the results to human-annotated sentiment analysis for more accurate findings. To address the study objectives, the following research questions are investigated:

1. What sentiment polarity is identified by both human evaluators and automated tools in social media comments on news articles evoking anger, fear, sadness and happiness?
2. What appraisal language structures represent each of the four basic emotions in the social media comments on the news articles?

The answers to the above questions provide insight into how sentiment, emotions and opinions can be extracted from text, both manually and automatically. This process aims to detect the emotional disposition of social media users of different ethnic backgrounds, which in turn helps decision makers to adjust policies and mitigate harmful effects of various events. The paper is organized into five main sections: section 1 provides the introduction comprising the background, research gap and research questions, section 2 reviews the related literature, section 3 describes the underlying methodology, section 4 presents the results, discussion, limitations and implications for future research, and section 5 finally wraps up with the conclusion.

2. Literature Review

The appraisal theory has often been examined in previous research on language sentiment as it focuses on how language conveys feelings and emotions towards people and phenomena (Wendland et al., 2018; Zaytoon, 2019; Zeng et al., 2024). The theory proposed by Martin and White (2005) builds on Halliday’s systemic functional grammar (Halliday, 1994) that divides language functions or meanings into ideational, interpersonal and textual. With its emphasis on the language of evaluation, the appraisal theory provides insight into how the interpersonal meaning is performed through the speakers’/writers’ expression of feelings and opinions towards people and phenomena.

The appraisal framework involves three domains: attitude, engagement and graduation. Attitude, which is investigated in the current work, has often been associated with the expression of feelings and opinions as it consists of three sub-systems: affect, judgment and appreciation, all of which serve to describe the interlocutor's emotions and views towards people and surroundings. Within the attitude domain, affect in language is described through positive or negative expressions showing emotional reaction. Judgment is made through positive or negative evaluation of human behavior or character in terms of social and conventional norms. Finally, appreciation is conveyed through the assessment of the value of objects, states of affairs or processes in terms of aesthetics or social valuation (Martin & White, 2005).

A number of studies applied the appraisal theory to analyze sentiment in online discourse. Wendland et al. (2018) conducted a sentiment analysis of Australian tweets on a hostage incident in 2014, using Martin and White's (2005) appraisal theory. They found that although the incident discussed in the tweets was expected to have a negative polarity, the tweets' analysis showed a number of positive emotions evoked by the situation. The positive emotion expressions detected included prayers and words encouraging solidarity, security and emotional support among community members in such a difficult situation.

In another study, Zaytoon (2019) used Martin and White's (2005) appraisal framework to analyze the sentiments expressed in Trip Advisor online reviews. She found that the analysis provided adequate description of the reviewers' emotions and viewpoints through studying the theory's sub-systems of judgment and appreciation. The reviews' analysis expressed the evaluative functions of judgment and appreciation more than affect since they reflected the reviewers' emotions towards humans (e.g. hotel staff) and phenomena (e.g. hotels, services, places, etc.) rather than the reviewers' own personal emotions.

Zeng et al. (2024) carried out a sentiment analysis of editorials related to the Russian-Ukraine conflict by creating a corpus annotation of the editorials' evaluative language according to Martin and White's (2005) appraisal scheme. Their work highlights a number of challenges in the annotation of evaluative language according to the appraisal theory, mainly due to the complexities involved in the recognition and categorization of evaluative expressions within the attitude sub-system.

However, they propose solutions to address these challenges, mainly through the annotation of chunks of expressions related to a particular emotion, with the aim of achieving more transparency and consistency in sentiment annotation of opinion texts (Zeng et al., 2024).

Sentiment analysis and emotion detection are sometimes used interchangeably, yet the main difference lies in the polarity aspect of sentiment analysis, aiming to label data as positive, negative or neutral, and the emphasis of emotion detection tools on determining the language user's emotional state or mood (Nandwani & Verma, 2021). The interdisciplinary study of sentiment analysis has recently been explored by an increasing number of researchers in the fields of psychology, linguistics and computer science (Peslak, 2018; Jaidka et al., 2018; Kausar et al., 2020; Ohiagu, 2020; Cahyanti et al., 2021; Kastrati et al., 2021; Boyd et al., 2022; Ibrahim et al., 2022; Date et al., 2023; Zeng et al., 2024). Most sentiment analysis studies use computer-assisted text analysis tools to detect sentiment polarity and emotional disposition of writers and/ or speakers, but there is limited research on human-conducted sentiment analysis. Several studies have found that automated sentiment analysis tools often produced less accurate results then human annotators due to machines' inability to detect context-related emotion words (Boukes et al., 2019; Jaidka et al.,2020; Gandy et al., 2025). Yet, no work, to the best of the author's knowledge, has compared human and automated sentiment analysis of online comments in situations triggering the basic human emotions of anger, fear, sadness and happiness, hence the contribution of the present study.

Most of the above-mentioned sentiment analysis studies were conducted using automated natural language processing classification tools such as Vader, Roberta, Naïve Bayes, SVM and LIWC. The author of the current study selected the LIWC tool, which is described in the following section, due to its ease of use, requiring no prior coding knowledge, alongside human annotation, to analyze the sentiment polarity and emotions within the data under study.

3. Methodology

The current study provides a quantitative-qualitative sentiment analysis of the comments made by Facebook users on news articles in order to detect sentiment polarity and emotion expression in these comments. The analysis of linguistic patterns indicating different emotions is based on Martin and White's (2005) appraisal theory, with emphasis on the function of attitude comprising the sub-functions of affect, judgment and appreciation.

The data collected is comprised of 712 comments, attached to 4 news articles, two of which including videos, triggering the four basic human emotions of anger, fear, sadness and happiness. The total number of words analyzed is 17600. For the analysis of the emotion of anger, a CNN article discussing Trump's proposal to evacuate Gaza for reconstruction, published in February 2025, was analyzed. For fear, an Al-Jazeera article covering the COVID-19 outbreak, published in April 2020, was selected. For sadness, an Arab News article about the drowning of a Moroccan boy in a well following failed rescue attempts, published in February 2022, was chosen. Lastly, for happiness, an Al-Jazeera article about the Moroccan football team's victory in the World Cup, published in December 2022, was selected. These four articles were selected on the basis of the topic's potential to trigger the four basic emotions as well as the number of comments (i.e. each had no less than 1000 words of comments).

The four selected articles align with the criteria established by psychological research to represent the basic emotions of anger, fear, sadness and happiness (Izard, 2010). The definitions of the basic emotions below are adapted from Izard (2010), who lays the foundation for expressing human emotions through various situational examples. Anger is defined as an emotion that results from the frustration of blocked goal responses. This emotion is likely to be caused by the news of the American president's plan to evacuate Gaza from its people, which causes frustration as the goal of achieving peace is blocked. The emotion of fear is defined as the sensation of harm from unpredictable circumstances leading to individuals' flight to one another for safety. This emotion is likely to result from the article about the outbreak of COVID-19 virus, which causes people to fear the unpredictable epidemic and search for safety. Sadness results from a life-changing loss, which is caused by the news about the loss of the Moroccan boy after days of rescue attempts.

Finally, happiness stems from the pride of achievement, which results from the news about the victory achieved by the Moroccan team in World Cup 2022.

The author chose to analyze comments attached to news articles as they usually contain more sentiment and emotional expressions than the news articles themselves (Anspach & Carlson, 2020). In order to collect the data, each of the four news articles was accessed on the newspaper Facebook page, and all the comments attached to it, appearing as ‘relevant comments’ were copied and pasted on a Word document. This was then followed by a data cleaning process in which irrelevant stop words, such as names, dates, emojis, special characters, non-English and non-meaningful words, were removed. Accordingly, 4 documents were prepared comprising comments on the articles evoking the 4 basic emotions. The total number of words in each of the documents representing comments on articles evoking anger, fear, sadness and happiness was 13480, 1053, 1161 and 1906, respectively. In order to obtain consistent values, regardless of the different data size, percentages of word frequencies were used in the analysis.

The data was then analyzed both quantitatively and qualitatively to identify linguistic patterns in sentiment and emotion expression. The quantitative analysis involved using the LIWC-22 sentiment analysis tool as well as human annotation. The LIWC-22 software was used to measure sentiment polarity, calculate positive and negative sentiment scores and detect prevalent emotions in each document. The human annotation was performed by two raters: the author and another holder of PhD in linguistics to manually calculate the number of positive and negative words in each document as well as the number of words indicating each of the four basic emotions. The average score of the two human raters was used to ensure consistency and unbiased measures. The automatic text analysis tool could provide quantitative measures of sentiment, but it could not capture context-dependent opinion words, which highlights the importance of human analysis alongside the automated tool. The human evaluators annotated the texts by assigning codes for sentiment attributes and then entered the texts into the AntConc corpus analysis tool to examine the relevant linguistic patterns in context.

4. Results and Discussion

This section presents the results of the data analysis performed using both automatic and human sentiment analysis methods. It aims to address the research questions proposed in section 1, by investigating the sentiment polarity of each set of comments, and the language structures representing the four basic emotions.

4.1 Sentiment Polarity

Sentiment polarity identifies the orientation of sentiment in a text, or, in other words, specifies whether a text has a dominant positive, negative or neutral sentiment. Table 1 below shows the sentiment polarity of the comments attached to the four articles evoking the 4 basic emotions of anger, fear, sadness and happiness. The sentiment polarity is calculated by LIWC-22 automatic sentiment analysis tool and shows that all the comments on the 3 articles representing topics triggering the negative emotions of anger, fear and sadness have more negative than positive emotions. However, although the tool measured more negative than positive emotion words in these comments, the overall tone, which expresses sentiment rather than emotion words (Boyd et al., 2022) was found to be more positive than negative in the comments attached to the topics evoking fear and sadness emotions. The comments on the article evoking positive emotions, on the other hand, had higher measures for both positive tone and positive emotions.

Article	**Emotion Evoked**	**tone_pos**	**tone_neg**	**emo_pos**	**emo_neg**
Gaza Evacuation	Anger	2.72	2.8	0.46	0.65
COVID-19 Breakout	Fear	2.87	2.6	0.36	0.44
Rayan’s Death	Sadness	5.48	2.93	0.37	1.87
Morocco’s Win	Happiness	11.78	0.59	5.7	0.23

Table 1: LIWC-calculated sentiment polarity of news article comments

In order to compare the measurements of the positive tone (tone_pos) versus positive emotions (emo_pos) on one hand, and the negative tone (tone_neg) versus negative emotions (emo_neg) on the other, the author analyzed the comments using LIWC’s option of coloring words representing each category. The colored words representing the positive tone included words of all grammatical categories bearing a positive connotation, even if the context in which they occurred conveyed a negative stance. For example, the comments on the article evoking anger included verbs like ‘clean’, ‘keep’, ‘play’, adjectives like ‘new’, nouns like ‘plenty’ and adverbs like ‘well’ as representative of a positive tone. However, these words were not chosen by the human annotators as indicating a positive emotion since their contexts were ‘attempts to *clean* Gaza’, ‘*keep* losing friends’, ‘*playing* diversion’, ‘*well*, they voted for him’, ‘a *new* casino’, and ‘*plenty* of lands that need attention’. Another emotional disposition that was not accurately detected by LIWC was the positive tone given to the comments on the article triggering fear. This was due to the tool incorrectly counting words as positive when they were not intended to convey positive emotions as in ‘government *playing* games, *super* misleading, and corona virus *won*’ and overlooking negative emotion words like ‘catastrophic, disaster, and misleading’.

The positive and negative emotion values detected by LIWC, on the other hand, included mainly adjectives such as ‘freaking, regretting, mad, and confused’ as negative emotion words and ‘happy, inspiring, courageous, funny, and proud’ as positive emotion words, yet the number of emotion words highlighted by LIWC was far less than the number assigned by the human annotators. This shows that automatic sentiment analysis tools may overlook a number of emotion words for favor of those included in their dictionaries. Table 2 below shows the difference between LIWC and human annotators in the emotion word count in percentages.

Article	**Emotion Evoked by commenters**	**LIWC-Measured Positive Emotion Words**	**Human-Measured Positive Emotion Words**	**LIWC-Measured Negative Emotion Words**	**Human-Measured Negative Emotion Words**
Gaza Evacuation	Anger	0.46	2.27	0.65	4.73
COVID-19 Outbreak	Fear	0.36	3.7	0.44	6.27
Rayan’s Death	Sadness	0.37	5.94	1.87	4.91
Morocco’s Win	Happiness	5.7	14.95	0.23	1.99

Table 2: Difference in emotion word counts between LIWC and human annotators

Table 2 reveals that although the human-measured emotion words were significantly more than those identified by LIWC, the latter could be generally effective in detecting positive or negative emotional disposition in a text. The automatic sentiment analysis tool assigned more negative emotion values to the comments attached to the 3 articles evoking anger, fear and sadness, and more positive emotion values to the comments on the fourth article triggering happiness. The same negative and positive emotional disposition was determined by the human annotators, but with more values given in the analysis of each text, and with more positive emotion values given by the human evaluators to the comments on the sadness-evoking article. This implies that automatic sentiment analysis tools like LIWC can be helpful in determining a text's sentiment or emotional disposition but through the measurement of positive and negative emotion words rather than positive and negative tone measurements, which were not very accurate as shown in Table 1.

A significant difference between LIWC and human measures is found in the values given to the emotion words related to the article evoking sadness, where human annotators assigned more positive than negative values for emotion words in the comments about that article. This is mainly due to the frequency of words referring to God's mercy and heaven, which are considered positive words and are usually used in cases of death to offer condolences and express sympathy. Yet, these words had higher counts in human measures than in LIWC measures, which again shows the automatic tool's tendency to ignore a significant number of emotion words either due to their absence from its dictionary or due to their association with other categories in the system like the category of religion.

Though LIWC is considered one of the most popular sentiment analysis tools, its measurements should be verified by human annotation to ensure more accurate results, which is one of the objectives of the current study. This is supported by previous research like Boukes et al. (2019), Jaidka et al. (2020) and Gandy et al. (2025). Boukes et al. (2019) studied the sentiment analysis of five automated instruments, including LIWC, aimed at detecting the sentiment disposition of Dutch news articles about the economy. The results showed that most of these tools yielded low-reliability scores compared to manual coding. Similarly, Jaidka et al. (2020) found that sentiment analysis tools like LIWC produced inconsistent results when analyzing English tweets for emotion words indicating the wellbeing of individuals in a number of US counties. The inconsistency was mainly due to differences in language use across individuals from varied geographical places and socioeconomic backgrounds, which could only be captured by human evaluators. In a third study, Gandy et al. (2025) compared the sentiment analysis of four NLP models, including LIWC, to human sentiment coding and found that these tools failed to detect the sentiment associated with the use of drugs in a selection of YouTube videos. They concluded that human coding remains to be the most reliable sentiment-detection method to date.

4.2 Appraisal language representation of the four basic emotions

This section analyzes the linguistic representation of basic emotions in the comments associated with the four news articles, followed by emotion word classification according to Martin and White's (2005) appraisal theory. The analysis of emotion words is based on both automatic and human ratings of emotion words in the collected texts.

The emotion of happiness is represented in LIWC as a positive emotion since the system does not distinguish between happiness-related words like ‘glad, satisfied, excitement’ and other positive adjectives not directly related to happiness like ‘easy, outgoing, brave, etc.’ For ease of comparison between human and automatic ratings, the human evaluators counted the positive emotion words in general, as shown in table 2 above. The values given to positive words in the happiness-evoking article about Morocco’s victory in World Cup were 5.7% by LIWC, and 14.95% by human raters. Yet, the human raters went one step further with the analysis of the positive emotions and counted the words related to happiness in all the comments on the above-mentioned article. They found that 6.87% of the words in the comments were related to happiness as in ‘proud, happy, glad, and achievement’. These values show that the most dominant emotion detected in the comments about the happiness-evoking article was actually that of happiness.

Table 3 below presents the frequencies (in percentages) of words associated with the three negative emotions of anger, fear, and sadness in the comments as identified by both LIWC and human annotators.

Article	Emot-ion	LIWC-Measured Anger	Human-Measured Anger	LIWC-Measured Fear	Human-Measured Fear	LIWC-Measured Sadness	Human-Measured Sadness
Gaza	**Anger**	**0.18**	2.28	**0.17**	0.29	**0.11**	0.14
COVID Breakout	**Fear**	**0**	2.18	**0.18**	2.46	**0.09**	0.66
Rayan's Death	**Sad-ness**	**0**	0	**0**	0	**1.57**	2.15
Morocco's Win	**Happi-ness**	**0**	0.73	**0**	0	**0**	0

Table 3 Frequencies of negative emotion words calculated by LIWC and human annotators

The table shows that in the comments on the article of Gaza's evacuation, likely to evoke anger, the most dominant negative emotion detected by both LIWC and human evaluators was anger (01.8 and 2.28). Yet, the human evaluators detected the anger emotion more accurately than LIWC since the latter rated a number of negative words in this article as fear-related though they were not actually associated with fear, which caused the values of anger-related and fear-related words detected by the tool to be close. For example, the automatic tool counted the different forms of the word 'stress', meaning emphasis, as fear-related words as in 'The British government stressed the need for …' and 'stressing that such plans represent an attack on the Palestinians.'

In the comments associated with the article evoking fear, related to the spread of COVID-19 pandemic, both LIWC and human evaluators assigned more values for the emotion of fear (0.18% and 2.46% respectively). However, the human raters also assigned a value to anger-related words in this article (2.18%), while LIWC did not assign any value for these words. This is attributed to the human raters' labeling of negative words like 'mayhem, misleading and lies' as anger-associated, which was not detected by LIWC in its word count of negative words.

The presence of anger-associated words in this article was mainly due to the commenters' condemnation of governments' policies in dealing with the pandemic.

The sentiment values calculated by both LIWC and human raters for the comments associated with the article about Rayan's death showed that the dominant emotion was sadness (1.57% and 2.15% respectively). This was mainly due to the absence of negative words associated with anger or fear in the discussion of a child's tragic death, which is likely to evoke readers' sympathy and sad sentiments.

In the article about Morocco's victory in World Cup 2022, which was likely to evoke happiness, no negative emotions were detected by LIWC, but a few negative emotion words were counted by the human raters (0.73%). These negative words mainly referred to the disappointment caused by the absence of other Arab winning teams and the fact that Morocco came in the third place rather than the first or the second.

In order to analyze sentiment expression in the news comments according to Martin and White's (2005) appraisal theory, the human raters classified the emotion words detected into the three categories of affect, judgment and appreciation. Words under the category of affect include words expressing human reactions like 'happy, glad, or sad'. Words describing judgment include words evaluating human behavior as in 'crazy, criminal, or deceitful.' Finally, words of appreciation assess the value of circumstances or processes as in 'heartbreaking experience, destructive policies, or ethnic cleansing'. Table 4 shows the percentage of emotion words belonging to each appraisal category in the news comments.

Article	**Emotion Evoked**	**Affect**	**Judgement**	**Appreciation**
Gaza Evacuation	Anger	0.65	2.4	1.63
COVID-19 Outbreak	Fear	0.57	2.37	3.32
Rayan's Death	Sadness	3.01	0.34	1.55
Morocco's Win	Happiness	2.83	10.65	1.46

Table 4 Emotion classification according to Martin and White's (2005) appraisal categories

Table 4 shows that in all comments, affect was dominant in the article evoking sadness, judgment was dominant in articles evoking anger and happiness, and appreciation was dominant in the article evoking fear. This implies that in commenting about the article evoking sadness, the commenters expressed their personal feelings of empathy and sorrow when talking about the death of Rayan, using words like ‘I’m very sad’ and ‘we’re broken-hearted’, but did not make judgments about people since it was a tragedy not caused by human intention. In the comments about the article of Gaza’s evacuation, likely to evoke anger, the dominant appraisal element was judgment since most commenters expressed their negative evaluation of the American president, Trump, who called for the controversial step of dislocating Palestinians. In the comments about COVID-19, the prevalent appraisal element was appreciation since most commenters evaluated the situation of the pandemic more than evaluating human behavior, though some comments also showed judgment of officials’ responses to the pandemic. In the comments about Morocco’s victory, likely to evoke happiness, the dominant appraisal element was judgment as the commenters evaluated the players, praising their talent and excellent performance.

4.3 Limitations and Recommendations for Future Research

The sentiment analysis of the news comments presented in this paper faces several limitations. The first limitation is the relatively limited corpus of comments related to the four news articles. The author chose the four articles as representative of the four basic emotions of anger, fear, sadness and happiness, but a larger corpus is recommended in future research to obtain a more comprehensive overview of sentiment representation in language. Another limitation is the lack of available information about the cultural background, age or gender of commenters. Future research could explore how culture, age and/or gender may influence sentiment polarity and emotion expression in language. A third limitation lies in the inaccuracies of some measures given by the LIWC-22 tool since it analyzes text across a large number of categories. The internal dictionary of LIWC includes thousands of words where each entry is part of a number of categories and sub-categories associated with different psychological constructs. For example, the word 'cried' is listed under different categories such as affect, positive tone, emotion, negative emotion, sadness emotion, verb, focus past, communication, linguistic and cognition (Boyd et al., 2022). Though the author selected only the categories related to affect when running the LIWC software, the existence of other semantic categories under which certain words are classified may have caused the automatic tool to overlook emotion-related meanings associated with those words. For example, the expression 'ethnic cleansing' may be absent as a negative emotion expression in LIWC's analysis due to its existence in another category like 'politics'. This suggests the need to revise the dictionaries of automated sentiment analysis tools and to include words in a wider context to achieve more accurate detection of emotion-related words.

5. Conclusion

This paper provides a sentiment analysis of news comments associated with articles likely to evoke the emotions of anger, fear, sadness and happiness. The analysis was conducted using both automatic annotation, carried out by the LIWC-22 tool, and human annotation. The results show that LIWC-22 assigned negative sentiment polarity for the comments associated with the articles evoking anger, fear and sadness, and positive polarity for the article evoking happiness. The same sentiment polarities were assigned by the human annotators but with higher values as they detected a larger number of emotion words than those detected by the automatic tool. Yet, the measures given by LIWC for positive and negative emotion words were more accurate in determining a text's sentiment polarity than the measures it gave for positive and negative tones since the latter incorrectly labeled a number of words as positive when in fact they were used in a negative context as in 'super misleading' and 'Corona virus won'.

The analysis of the appraisal language structures used in the news comments revealed that affect (language expressing human reactions) was dominant in the comments associated with sadness, judgment (language expressing human behavior) was dominant in the comments associated with anger and happiness, and appreciation (language evaluating events) was dominant in the comments associated with fear. The implications suggest that the appraisal framework could be used in sentiment analysis tools to detect the dominant emotional tone of a text, revealing whether it primarily reflects personal reactions, attitudes toward people, or attitudes toward events.

The present study proposes methods for improving the accuracy of sentiment analysis tools, primarily by incorporating cross-checking with human evaluators and refining automated sentiment analysis dictionaries to include more comprehensive contextual information for enhanced emotion detection. Accordingly, the findings of this study can guide the refinement of sentiment analysis tools and contribute to more accurate detection of emotional expression in future research.

References

Al Jazeera. (2020, April 20). *[Post linking to article titled "Italy sees first fall of active coronavirus cases: Live updates"]. Facebook.* https://www.aljazeera.com/news/2020/4/20/italy-sees-first-fall-of-active-coronavirus-cases-live-updates

Al Jazeera. (2022, December 21). *[Post linking to video titled "Morocco team arrives home after World Cup success"]* [Video]. *Facebook.* https://www.facebook.com/watch/?v=1402754950525206

Anspach, N.M. & Carlson, T.N. (2020). What to believe? Social media commentary and belief in misinformation. *Political Behavior, 42*, 697–718. https://doi.org/10.1007/s11109-018-9515-z

Arab News. (2022, February 5). *[Post linking to article titled "Tragic end for bid to save 5-year-old Rayan, trapped 30 meters down well in Morocco"]. Facebook.* https://www.facebook.com/ArabNews/posts/pfbid02nHnt7dya4WiQpHxjZeSnsUzcVqpA5wxx474R3h8cgEuooVGkzgEfYtkpp2MmzUiKl

Boukes, M. et al. (2019). What's the tone? Easy doesn't do it: Analyzing performance and agreement between off-the-shelf sentiment analysis tools. *Communication Methods and Measures*, *14*(2), 83–104. https://doi.org/10.1080/19312458.2019.1671966

Boyd, R. L. et al. (2022). The development and psychometric properties of LIWC-22. Austin, TX: University of Texas at Austin. https://www.liwc.app https://doi.org/10.13140/RG.2.2.23890.43205

Cahyanti, A. et al. (2021). Comparing the language style used by native and non-native English speakers in The Ellen Show. *English Education Journal*, *11*(4), 579-588. https://doi.org/10.15294/eej.v11i1.50290

CNN. (2025, February 5). *[Post linking to video titled "President Trump said the US 'will take over' Gaza, possibly using US troops, and described his vision for Gaza as a new 'Riviera'"]* [Video].

Facebook. https://www.facebook.com/cnn/videos/1337488610771302

Date, S. et al. (2023). Sentiment analysis using computer-assisted text analysis tools. *Proceedings of the International Conference on Applications of Machine Intelligence and Data Analytics (ICAMIDA 2022)*, *105*, 671-679. https://doi.org/10.2991/978-94-6463-136-4_58

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, *6*(3–4), 169–200. https://doi.org/10.1080/02699939208411068

Gandy, L.M. et al. (2025). Public health discussions on social media: Evaluating automated sentiment analysis methods. *JMIR Formative Research*, *9*, 57395. https://doi.org/10.2196/57395

Gu, S. et al. (2019). A model for basic emotions using observations of behavior in Drosophila. *Frontiers in Psychology, 10.* https://doi.org/10.3389/fpsyg.2019.00781

Halliday, M.A.K. (1994). *An introduction to functional grammar*. Oxford, UK: Oxford University Press.

Ibrahim F.A. et al. (2022). COVID19 outbreak: A hierarchical framework for user sentiment analysis. *Computers, Materials and Continua, 70* (2), 2507-2524. https://doi.org/10.32604/cmc.2022.018131

Izard, C. (2010). The many meanings/aspects of emotion : Definitions, functions, activation, and regulation. Emotion Review, 2 (4), 363-370. https://doi.org/10.1177/1754073910374661

Jack R. et al. (2014). Dynamic facial expressions of emotion transmit an evolving hierarchy of signals over time. Current Biology, 24,187–192. https://doi.org/10.1016/j.cub.2013.11.064

Jaidka, K. et al. (2018). Predicting elections from social media: a three-country, three-method comparative study. *Asian Journal of Communication*, *29*(3), 252–273. https://doi.org/10.1080/01292986.2018.1453849

Jaidka, K. et al. (2020). Estimating geographic subjective well-being from Twitter: A comparison of dictionary and data-driven language methods. *Proceedings of the National Academy of Sciences, 117*. 201906364. https://doi.org/10.1073/pnas.1906364117

Kastrati, Z. et al. (2021). A deep learning sentiment analyser for social media comments in low-resource languages. *Electronics, 10* (10). https://doi.org/10.3390/electronics10101133

Kausar, S. et al. (2020). A sentiment polarity categorization technique for online product reviews. *IEEE Access, 8,* 3594-3605. https://doi.org/10.1109/ACCESS.2019.2963020

Martin, J.R. & White, P.R.R. (2005). *The language of evaluation: Appraisal in English.* New York: Palgrave, Macmillan.

Nandwani, P., & Verma, R. (2021). A review on sentiment analysis and emotion detection from text. *Social network analysis and mining*, *11*(1), 81. https://doi.org/10.1007/s13278-021-00776-6 https://doi.org/10.1097/EDE.0000000000001671

Ohiagu, O.P. (2020). Variations of English language use on Facebook by select native and non-native Speaker. In: Giri, R.A. et al. (eds). Functional variations in English. *Multilingual Education*, *37.* Cham: Springer. https://doi.org/10.1007/978-3-030-52225-4_10

Peslak, A. (2018). Facebook fanatics: A linguistic and sentiment analysis of the most “fanned” Facebook pages. *Journal of Information Systems Applied Research*, *11*(1), 23-33. http://jisar.org/2018-11/ ISSN: 1946-1836

Wendland, J. et al. (2018). Sydney Siege, December 2014: A visualisation of a semantic social media sentiment analysis. ISCRAM 2018 Conference Proceedings – 15th International Conference on Information Systems for Crisis Response and Management. Rochester, New York. https://www.researchgate.net/publication/325258793_Sydney_SiegeDecember_2014_A_Visualisation_of_a_Semantic_Social_Media_ Sentiment_Analysis

Wilson-Mendenhall et al. (2013). Neural evidence that human emotions share core affective properties. *Psychological Science, 24* (6), 947-956. https://doi.org/10.1177/0956797612464242

Zaytoon, H. (2019). Sentiment analysis: An appraisal approach to Trip Advisor reviews headlines. *Journal of scientific research in arts, 9,* 207-227. https://doi.org/10.21608/jssa.2019.75640

Zeng, J. et al. (2024). Annotating evaluative language: Challenges and solutions in applying appraisal theory. In H. Bunt (Ed.), Proceedings of the 20th Joint ACL - ISO Workshop on Interoperable Semantic Annotation, 144-151. *European Language Resources Association (ELRA).* https://aclanthology.org/2024.isa-1.17