




# Arab reactions towards Russo-Ukrainian war

Moayadeldin Tamer<sup>1</sup>, Mohamed A. Khamis<sup>1,2\*</sup> , Abdallah Yahia<sup>1</sup>, SeifALdin Khaled<sup>1</sup>,  
Abdelrahman Ashraf<sup>1</sup> and Walid Gomaa<sup>1,3</sup>

\*Correspondence:

[mohamed.khamis@ejust.edu.eg](mailto:mohamed.khamis@ejust.edu.eg);  
[maomar@ejada.com](mailto:maomar@ejada.com)

<sup>1</sup>Egypt Japan University of Science  
and Technology, P.O. Box 179,  
21934, New Borg El-Arab City,  
Alexandria, Egypt

<sup>2</sup>Ejada Systems Ltd., 620-622 El  
Horeya Street, Zizinia, Alexandria,  
Egypt

Full list of author information is  
available at the end of the article

## Abstract

The aim of this paper is to analyze the Arab peoples reactions and attitudes towards the Russo-Ukraine War through the social media of posted tweets, as a fast means to express opinions. We scrapped over 3 million tweets using some keywords that are related to the war and performed sentiment, emotion, and partiality analyses. For sentiment analysis, we employed a voting technique of several pre-trained Arabic language foundational models. For emotion analysis, we utilized a pre-constructed emotion lexicon. The partiality is analyzed through classifying tweets as being 'Pro-Russia', 'Pro-Ukraine', or 'Neither'; and it indicates the bias or empathy towards either of the conflicting parties. This was achieved by constructing a weighted lexicon of n-grams related to either side. We found that the majority of the tweets carried 'Negative' sentiment. Emotions were not that obvious with a lot of tweets carrying 'Mixed Feelings'. The more decisive tweets conveyed either 'Joy' or 'Anger' emotions. This may be attributed to celebrating victory ('Joy') or complaining from destruction ('Anger'). Finally, for partiality analysis, the amount of tweets classified as being 'Pro-Ukraine' was slightly greater than Pro-Russia' at the beginning of the war (specifically from Feb 2022 till April 2022) then slowly began to decrease until they nearly converged at the start of June 2022 with a shift happening in the empathy towards Russia in August 2022. Our Interpretation for that is with the initial Russian fierce and surprise attack at the beginning and the amount of refugees who escaped to neighboring countries, Ukraine gained much empathy. However, by April 2022, Russian intensity has been decreased and with heavy sanctions the U.S. and West have applied on Russia, Russia has begun to gain such empathy with decrease on the Ukrainian side.

**Keywords:** Natural language processing; Arabic text analysis; Sentiment analysis; Emotion analysis; Partiality analysis; Twitter; Deep learning

## 1 Introduction

The Ukrainian crisis is one of the most complicated and unfortunate events of this decade with many aspects to be considered to have an informed opinion about. Social media platforms (e.g., Facebook, Twitter, etc.) are currently the main data source for public opinion analysis [1]. A great deal of work has been done to analyze public opinion on ongoing affairs and to study the influence of such events on people. For instance, the authors in [2] showed that the 2012 Olympic Summer games, held in London, increased the life satisfaction and happiness of Londoners during the Olympics period, particularly around the

© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

opening and closing ceremonies. There were no consistent changes (either positive or negative) in anxiety during this period in comparison to residents in neighboring cities such as Paris and Berlin.

Amid the Brexit controversy, the researchers in [3] studied the public attitudes towards the EU (European Union) testing the effect of “real world” arguments on both sides of the campaign that attempted to influence the vote through pro-EU or anti-EU messages. Their main finding was that the pro-EU arguments had the potential to significantly increase the support for “remaining” in the union whilst the anti-EU arguments had less potential to impact the support for either “remaining” or “leaving”.

As mentioned in [4] Twitter data are an important source for studying public response, and thus utilized to examine COVID-19 related discussions, concerns, and sentiments emerged from tweets. The results indicated that the dominant sentiment for the spread of coronavirus was “anticipation” followed by mixed feelings of “trust”, “anger”, and “fear” for different topics and significant feelings of “fear” when new cases and deaths were discussed. In [5], the authors presented a Twitter dataset of the Russo-Ukrainian war. The majority of the tweets are written in the English language (about 60%). Until the day their paper was written (7th of April, 2022) the dataset has reached 57.3 million tweets written by 7.7 million users.

In [6], the authors also provided a Twitter dataset of the Russo-Ukrainian conflict. The data collection process was not filtered by any language or geographical location. Thus, the dataset includes tweets in several languages from different regions. The authors did some descriptive analysis over this dataset. For example, an analysis for the daily volume of the tweets revealed that an average of about 200,000 tweets have been posted daily. The authors also presented the number of tweets containing the keywords used in data crawling; this revealed that most of the tweets contained the keyword ‘putin’ (328,186 tweets) followed by ‘zelensky’ (86,122 tweets). Moreover, the authors also presented the top-10 used hashtags and mentions. This analysis reveals that Zelensky had the highest mentions, followed by NATO and other western leaders. A word cloud of tweet text was also provided that showed the significance of tokens like ‘breaking’, ‘news’, and ‘suspensions’.

The main interest of the current work is detecting and analyzing how people in the Arabic-speaking Middle and Near East reacted towards the *Russo-Ukraine War* and its related parties as the conflict unfolded. It is true that the Middle/Near East, and the Arab World in particular, are not directly involved in this conflict, but it has direct and indirect dire consequences such as the case with oil prices, food and other vital commodities and goods; and the pressure from either side of the conflict to attract support from the region. As mentioned in [7], African economies – including Arab African countries – have been the worst losers aside from Ukraine, should this war escalate further ahead. While the region’s oil, gas, and commodity-exporting countries have benefited from the rising energy and commodity prices, strong negative effects have befallen upon Other Arabic as well as African countries with their huge dependence on Russian and Ukrainian food imports and other essential metal and oil products.

The different reactions towards the conflict are apparent and in order to better understand them we aim to know the public’s perception towards the war, and to better understand whether people are in favor of which side and to what extent. The approach we followed to achieve this was the collection of social media posts, particularly tweets, that are related to the war incidents with the required preprocessing achieved. Specifically, our

**Table 1** Search queries in our tweets dataset scraping procedure

Query No.	Query Text	English Translation	Query No.	Query Text	English Translation
1	روسيا أوكرانيا	Russia Ukraine	21	قوات روسيا	Russia Forces
2	بوتين	Putin	22	العملية العسكرية أوكرانيا	Ukraine military operation
3	كييف	Kyiv	23	العملية العسكرية الروسية	Russian military operation
4	#الحرب_الروسية_الأوكرانية	The Russian Ukrainian War	24	دونباس	Donbass
5	الناتو	NATO	25	هجوم روسيا	Russia attack
6	الرئيس الأوكراني	Ukrainian President	26	دونيتسك	Donetsk
7	القوات الروسية	Russian forces	27	أوديسا	Odessa
8	#الحرب_العالمية_الثالثة	Third World War	28	قصف أوكرانيا	Ukraine bombing
9	الجيش الروسي	Russian Army	29	ماريوبول	Mariupol
10	الجيش الأوكراني	Ukrainian army	30	القصف الروسي	Russian bombing
11	القوات الأوكرانية	Ukrainian forces	31	سومي	Sumy
12	الغزو الروسي	Russian invasion	32	خطر أوكرانيا	Ukraine risk
13	زيلينسكي/زيلنسكي	Zelensky	33	ضرب أوكرانيا	Ukraine hit
14	العاصمة الأوكرانية	Ukrainian capital	34	#WorldWar3	#WorldWar3
15	غزو أوكرانيا	Ukraine invasion	35	بحر ازوف	Sea of Azov
16	بيلاروسيا	Belarus	36	لوهانسك	Luhansk
17	نووي/نووية	Nuclear	37	الاتحاد السوفيتي	Soviet Union
18	خاركيف	Kharkiv	38	مجلس الامن	Security Council
19	الهجوم الروسي	Russian attack	39	الشعب الأوكراني	Ukrainian people
20	قوات أوكرانيا	Ukraine forces	40	الرئيس الروسي	Russian President

tweets scraping has started from the 23rd of February 2022 (the day before the start of the war) till the 31st of Jan 2023. Data scraping was based on search queries using Twitter trends of that period (e.g., الحرب العالمية الثالثة en: World War 3), names of personnel and institutions being mentioned frequently during the war (e.g., بوتين و الناتو en: Putin and NATO) or places witnessing the war (e.g., كييف en: Kiev). The total number of queries used in the search were 40 different queries incorporating uni-grams, bi-grams, and tri-grams. Table 1 presents the exact search queries in the data scraping procedure for the tweets in our dataset.

Social media posts compared with survey polls can result in better and more thorough perception of public opinion about specific topics in a better scientific manner [1]. Moreover, social media (such as Twitter) currently plays a major role in affecting the public opinion and attitude as has been observed by several studies [8]. Results of this study showed how an election candidate in the U.S. can influence other users to change the course of the election by identifying high in-degree centrality within users participating in a political discussion as happened in the 2012 and 2016 U.S. presidential elections. The authors in [9] concluded that automated *public opinion monitoring* using social media is a very powerful tool, able to provide interested parties with valuable insights for more fruitful decision making. Twitter has been gaining significant attention in this respect, since people use it to express their views and politicians use it to reach their voters, in a very short, concise, yet effective way.

In the current work we have performed three different kinds of analyses over the collected Twitter data. The first is concerned with the *sentiment*: what the people's attitudes have been towards the war expressed through tweeting short condensed text. This can

be either positive, negative, or neutral. The second kind of analysis is concerned with analyzing *emotion* with six different emotions. We adopted a *lexicon*-based approach for emotion analysis. This approach calculates the semantic orientation of specific text (e.g., documents, tweets, posts, etc.) from the semantic orientation of its lexicon (by aggregating the scores of the individual  $n$ -grams in the text). The learning approach is a bit harder at this stage as emotion datasets for the Arabic language are very scarce and very limited in size. The final analysis is concerned with the *people's bias/partiality analysis* towards either of the two parties directly involved in the conflict. This gives an indication of the credibility and propaganda success of each of the conflicting parties, at least throughout the Arab region.

Performing several kinds of analyses aims at detecting the impressions and opinions expressed in different forms. More specifically, *sentiment analysis* mainly aims to express the main attitude behind the tweet, *emotion analysis* searches for specific strong feelings in the tweet, while *partiality analysis* mainly aims to know how people favor either side, which in the current geopolitical context can convey significant trends of the public opinion that may have an impact on the decision makers towards the current and future international situation. It is important to note that our assumption for being 'Pro-Russia' involves favoring Russia and/or its supposed allies, disapproving the narrative of the U.S., Ukraine, and/or the west. On the other hand, standing by the war while being 'Pro-Ukraine' involves favoring Ukraine, U.S., or the west, disapproving Russia and/or its supposed allies, and disapproving the war and/or the Russian's narrative.

The paper is organized as follows. Section 1 is an introduction. Section 2 presents a background about the techniques and approaches utilized in the proposed analysis. In addition, this section presents some related works that utilize Twitter data to analyze how people react towards specific topics and events. Section 3 presents our methodology including the data collection process and the three types of analyses performed over the data. Section 4 presents the experimental work performed to analyze the reactions with respect to the three different aspects along with the results and discussions. Finally, Sect. 6 concludes the paper with pointers to future work.

## 2 Background and related works

Natural Language Processing (NLP) has significantly increased its potential in the last several years, with the newly trained foundational large models and their impacts on NLP applications. NLP is a subfield of computer science, artificial intelligence, and linguistics that is concerned with developing computational tools to understand text and speech in a similar way to humans. This includes the interactions between computing devices and humans as well as programming computers to process and analyze big chunks of human language. Arabic NLP is the application of NLP tools and technologies, particularly artificial intelligence and text mining, to understand the Arabic language in general; and particularly, Modern Standard Arabic (MSA) and the different Arabic dialects. These include, Arabic text search, PoS (Part-of-Speech) tagging, translation, diacritization, sentiment and emotion analyses, topic modeling, document summarization, etc.

Our focus in this article is to apply Arabic NLP techniques to extract semantic insights from Twitter text data.

## 2.1 Sentiment analysis

Sentiment analysis, among other approaches, represents a decent percentage of Arabic applications which have led to impressive discoveries. In [10] the authors mentioned different approaches to Twitter sentiment analysis including machine learning, lexicon-based, and hybrid-based approaches. In our work we utilized machine learning pre-trained models to perform sentiment analysis. We determine the sentiment of a tweet by taking the majority voting of three of the most well known state-of-the-art models that are used to predict sentiment in Arabic: (1) the Mazajak model [11] built on a Convolutional Neural Network (CNN) [12] followed by a Long-Short Term Memory (LSTM) [13], (2) AraBERT [14], a transformer-based model inspired by the Google BERT model [15], (3) CaMeL-Tools [16] whose driving design principles were largely inspired by the MADAMIRA [17], Farasa [18], CoreNLP [19], NLTK [16]. AraBERT can be fine-tuned on different datasets; we decided to fine-tune it on ArSAS [20] a multi-class dataset where sentences are classified under one of the following classes: 'Positive', 'Negative', 'Neutral', or 'Mixed'.

Mazajak [11] is built on a Convolutional Neural Network (CNN) followed by a Long-Short Term Memory (LSTM). The word embeddings of Mazajak were built from a corpus of 250 million different Arabic tweets. The tweets were scrapped through time periods between 2013 and 2016. LSTM is a recurrent neural network (RNN) used heavily in AI and deep learning applications, specially for modeling and analyzing sequential data. Unlike standard feedforward neural networks, LSTM has feedback connections to handle sequential data and realizes memorization for that purpose. This recurrent network can process single data points (e.g., images) as well as entire sequences of data (e.g., text, speech, video, etc.). The notion of LSTM stems from the analogy that a typical RNN should have both "long-term memory" and "short-term memory". The weights and biases in the network connections change once per epoch of training. This is analogous to how physiological changes in synaptic strengths that store long-term memories. The activation patterns in the network change once per time-step.

The second pre-trained model is AraBERT [14] which is a transformer-based model inspired by the Google BERT model [15]. AraBERT is based on manually extracted Arabic news websites. The authors used two publicly available large Arabic corpora: Arabic Corpus [21] consisting of 1.5 billion words including more than 5 million articles collected from 10 main news sources covering 8 countries, and the Open Source International Arabic News Corpus OSIAN [22] consisting of 3.5 million articles (~1B tokens) extracted from 31 news sources covering 24 Arab countries; the pre-training dataset final size is 70 million sentences (after duplicate sentences were removed), that is ~ 24GB of text. The sentiment voting model AraBERT [14] is an Arabic language model that is based on BERT. Bidirectional Encoder Representations from Transformers (BERT) is a transformer-based ML model for NLP pre-training developed by Google [15]. The original English-language BERT has two components: (1) the BERT-BASE which consists of 12 encoders with 12 bidirectional self-attention heads, and (2) the BERT-LARGE which consists of 24 encoders with 16 bidirectional self-attention heads. Both models are pre-trained from unlabeled data extracted from the BooksCorpus with 800M words and English Wikipedia with 2500M words. The last tool used to vote for sentiment analysis is CaMeL-Tools [16] whose driving principles of the design were largely inspired by the MADAMIRA [17], Farasa [18], CoreNLP [19], and NLTK [16].

The authors in [23] presented a linguistically accurate, large-scale morphological analyzer for the Egyptian Arabic language, which differs from the Modern Standard Arabic (MSA) phonologically, morphologically, and lexically and has no standardized orthography. The authors in [24] presented ADIDA, a system for automatic dialect identification for Arabic text distinguishing between dialects of 25 Arab cities in addition to the MSA. A Dialect Identification system in CaMeL-Tools [16] was used as the back-end component of ADIDA for computing the dialect probabilities of the given input. Arabic has a distinguishing characteristic of its complex structure that a computational system has to deal with at each linguistic level [25]. Beside the structure, one of the most challenging difficulties is how the analysis may differ with different Arabic dialects.

Twitter sentiment analysis is concerned with analyzing users' tweets in terms of thoughts and opinions in a variety of domains. This analysis can be very important for researchers who need to understand people's views about a particular topic or event [10].

In [26] the authors collected a dataset that was clustered into three categories highlighting the use of social media in the 'Jammu and Kashmir' conflict and identified how people utilized Twitter to reach many others at the same time without any geographical restrictions to raise awareness of the situation in Kashmir by using hashtags, retweets, or replies to tweets. In [27] the authors collected 43,000+ tweets of Donald Trump, trying to identify patterns in his tweets and identify changes over time and how entering politics has affected his behavior on social media. Also, identifying topics that the former 45th president of the U.S. discussed on Twitter.

In [28] the authors collected 1,433,032 tweets, extracting 57,842 tweets filtered by Hurricane Florence in 2018 between August and October. Their analysis showed that human sentiment plays an important role in spreading disaster information compared to the news of the hurricane in online communication. Moreover, people actively utilized Twitter to share a lot of emotions, opinions, and information about the Hurricane; concluding that governments and decision makers should monitor Twitter data to understand the human environment.

In [29] the authors collected a sentiment-annotated dataset for the analysis of Brazilian protests in 2013 annotated by three raters. Each document was classified in one of three classes: positive, negative, or neutral with 56% being classified as Neutral and only 4% as Negative.

Regarding work related to the Arab world and the Arab language, in [30] the authors tried to understand the roots of ISIS terrorist group and its supporters' using data collected from Twitter classifying tweets to "Pro-ISIS" and "Anti-ISIS", and then going back to analyze the historical timelines of both kinds of users supporting and opposing, looking at their pre-ISIS period. One of the conclusions reached was that ISIS supporters refer a lot more to the Arab Spring uprisings that failed.

In [31] the authors aimed to predict online Islamophobic behavior after the Paris terrorist attacks on the 13th of November 2015, through collecting millions of tweets related to these attacks. Tweets are then identified mentioning Islam and Muslims going through attitudes towards Islam and Muslims before the attack. The authors built a classifier to predict post-event stance towards Muslims utilizing pre-event interactions.

In [32] the authors investigate the emotional intensity of students' public opinion on the Internet. The authors studied the challenges of feature selection in sentiment tendency analysis. Sentiment analysis of students' cross-media written text is done through an im-



proved MapReduce combinator model. In [33] the authors utilized the Guardian newspaper and extracted useful information about the world points-of-view on the important events in Egypt, from early 2011 onwards, to detect the world perception for such events. The authors did sentiment analysis on the articles included in the 'World' section spanning the period from the start of 2010 to the end of 2017 using the 'Egypt' keyword. The authors got the uni-gram tokens from every article and utilized these tokens to infer the sentiment using three lexicon dictionaries: afinn, nrc, and Bing. The analytics indicated that the common trend was slightly negative during the whole selected period. Some conflicting feelings were appearing during this time span e.g., positive, negative, trust, fear, anger, and anticipation. The findings showed also that the years 2011 and 2013 had the peaks in both of positive and negative sentiments attributed to the two uprisings in Egypt.

## 2.2 Emotion analysis

Emotion analysis is the process of identifying and analyzing the underlying emotions expressed in text (e.g., [32, 34–41]). Emotions, in our context, can be mainly one of six classes: anger, disgust, fear, joy, sadness, and surprise. Our emotion analysis adopts a *lexicon-based approach*.

The lexicon-based approach is one of the main approaches for semantic analysis; it measures the semantic class of a specific text from the semantic orientation of its words [42]. The semantic class can be positive, neutral, or negative [42] or emotion class (e.g., anger, disgust, fear, joy, sadness). Specifically, the lexicon-based approach uses a semantic lexicon to score a document by aggregating the semantic scores (or taking the majority class(es)) of all the words in this document. The semantic lexicon contains a word and its corresponding semantic score [43].

Most works of opinion analysis in English can depend successfully on sentiment lexicon like SentiWordNet, e.g., [44–47]. However, Arabic sentiment lexicon faces some challenges, e.g., limited size, usability issues considering the Arabic rich morphology, public unavailability, and the huge diversities among the different dialects. The authors in [48] addressed these issues and created a publicly available large scale Standard Arabic sentiment lexicon (ArSenL) using a combination of existing resources: English SentiWordNet, Arabic WordNet, and the Standard Arabic Morphological Analyzer (SAMA). The authors evaluated their proposal in terms of subjectivity and sentiment analysis.

In [49] the authors presented a way to build an electronic Arabic lexicon by using a hash function that converts each word as an input to a corresponding unique integer number being used then as a lexicon entry. In [50] a large-scale sentiment lexicon called MoAr-Lex was presented; it was built through a novel technique for automatically expanding an Arabic sentiment lexicon using word embedding. The authors evaluated the quality of the automatically added terms in multiple ways. One of the advantages is its ability to incorporate terms that are commonly used in social media, but would normally be considered misspelled such as جميل (beautiful) with the last Arabic letter wrongly repeated twice.

In [51], the authors showed that the use of a sentiment lexicon (whether scored or not) has improved the sentiment classification results while the use of the scored lexicon consistently showed best classification results. Their experiments also showed that the use of scored lexicon can increase the sentiment classifier's ability to generalize across multiple datasets.

In [52] the authors presented AraVec, a collection of pre-trained Arabic word embedding models that can be used for Arabic NLP tasks (e.g., sentiment analysis, emotion analysis, etc.). AraVec is an open source and free to use project. The first version of AraVec contains six word embedding models. These models are built on top of three Arabic content channels: Twitter, World Wide Web pages, and Wikipedia Arabic articles. The total number of tokens used to build the models is more than 3,300,000,000.

In [34] the authors presented a deep learning approach for multi-label emotion classification of Arabic tweets. The proposed model is a multilayer Bidirectional-Long Short Term Memory (BiLSTM) trained on top of pre-trained word embedding vectors using the SemEval2018 Task1 dataset [53]. Several pre-processing steps are applied, e.g., normalization, stemming, replacing the most common emojis with their meanings using a manually constructed emoji lexicon. Word embedding was found to be the best method for feature generation. The AraVec [52] pre-trained word embedding model with Continuous Bag of Words (CBoW) avails 300 dimensional word vectors for each word in the dataset [53]. The average embedded word vector is then calculated for each tweet, then the BiLSTM is used for classification. The proposed method achieved the best results compared with Support Vector Machines (SVM), Random Forests (RF), and the fully connected deep Neural Network (DNN). It achieved 9% increase in the validation results compared to the previously best obtained results by SVM.

In [35] the authors provided a practical overview on developing an Arabic language model for emotion classification of Arabic tweets. In [36] the authors classified emotions in Arabic tweets: joy, anger, sadness, and fear. The proposed model is based on a deep Convolutional Neural Network (CNN) and word vectors trained specifically on the used dataset. The proposed deep learning approach was evaluated on the Arabic tweets dataset provided by SemiEval for the EI-oc task [53]. The model achieved high training accuracy of 99.90% and validation accuracy of 99.82%. The authors compared their results with three other ML approaches: SVM, Naïve Bayes (NB), and Multi-Layer Perceptron (MLP); implemented using three different Arabic stemmers (Light stemmer, ISRI, and Snowball), and two basic feature extractors (word count and TF-IDF).

In [38] the authors proposed a Bayesian inference method for emotion analysis in different semantic dimensions and inferred the co-occurrence of multiple emotion labels from the words in the document. The experiment is performed on the Chinese emotion corpus, i.e., Ren-CECps [54] which has high accuracy and is robust in word and document emotion predictions.

In [39] the authors used hashtags to label emotions. The method was evaluated by two subject studies: through psychology experts and through general crowd. The labels generated by experts were consistent with the hashtag labels of Twitter messages in more than 87% of the cases. The authors developed Emotex which is a supervised learning approach that classifies Twitter messages by the emotion classes they represent. Emotex correctly classifies the emotions presented in more than 90% of the text messages.

In [40] the authors studied various ML-based methods for emotion detection. The methods include ANN and DL. The ANN approaches were the Perceptron and Multilayer Perceptron. The DL approaches were the CNN-LSTM, CNN-BiLSTM, CNN-GRU, CNN-BiGRU, BiLSTM, and CNN. The authors used various feature representation approaches like n-grams, TF-IDF, word-embeddings, and contextualized embeddings. The authors evaluated the algorithms on the “International Survey on Emotion Antecedents and Re-



actions” (ISEAR) dataset [55]. The results showed that the model consisting of BERT with dense layer outperformed all other methods with macro-average F1-measure equals 0.71 for seven emotions, 0.76 for five emotions, and 0.8 for four emotions.

In [56] the author studied the Arabic songs and lyrics of the very famous singer Abd ElHalim Hafez (عبد الحليم حافظ). The work of the artist has many varieties with a big range of genres spanning romanticism, nationalism, spiritualism, etc. The author analyzed the common characteristics of the artist’s work comprising the composers and lyricists that the artist had been working with. The same author in [57] studied the lexical density and diversity of the same singer Abd ElHalim Hafez (عبد الحليم حافظ). The author analyzed the most important words, idioms, and tokens performed in the songs using word clouds and term frequency-inverse document frequency (TF-IDF). The author had shown a tight correlation between the analysis statistically and the political and social status in Egypt and the Arab region at that time. The author also studied the effectiveness of Part-of-Speech (PoS) tagging in genre analysis and classification.

### 2.3 Partiality analysis

Our final analysis is concerned with the people’s bias/partiality analysis towards either of the two parties involved in the conflict. Partiality analysis mainly aims to know how people favor either side. This analysis can convey significant trends of the public opinion that may have an impact on the decision makers. This analysis may be the most challenging with attempting to determine the amount of empathy each party receives across the data.

For partiality analysis, we first filter out the tweets with ‘Neutral’ or ‘Unspecified’ sentiment, as these tweets are expected not to carry any bias towards the conflicting parties. The resulting dataset then contained about 0.4M tweets that are used to pretrain an LSTM-based Neural Network. The aim is to classify tweets of the dataset as being either: ‘Pro-Russia’, ‘Pro-Ukraine’, or ‘Neither’.

In [58] the authors studied the effect of feature selection metrics on the performance of Decision Trees, Naïve Bayes classifiers, and Support Vector Machines. The evaluation is done through bias analysis of highly skewed data. Three types of biases are metric bias, class bias, and classifier bias. Experiments were performed to study the employment of these biases together in an efficient way to achieve good classification performance. The authors reported the results and best methods for text classification based on bias analysis. Over-sampling is found to be an effective way for class bias handling. In [59] the authors analyzed real instances of manual edits aimed to remove bias from Wikipedia pages. In [60] the author analyzed the partiality in Italian translations of three articles on Italian politics published in 2015 in the New York Times and the Financial Times. It looks at the discursive re-localization of these three translations when being distributed in the form of Italy’s politics and media.

In [61] the authors proposed a statistical model to identify biased users and social bots sharing the biased Twitter content. The authors used annotated twitter dataset and checked the results of sentiment analysis with and without the biased tweets and studied the biased users effects at micro-level and macro level. The results showed that the proposed approach is effective in identifying the biased users and bots from other authentic users using sentiment analysis. In [62] the authors used Twitter data from the 2018 U.S. midterm elections. The authors proposed a method to detect voters on Twitter and compare their behaviors with various accounts sampled randomly. Some accounts flood the

public data stream with political content sinking the voters' majority vote. Consequently, these hyperactive accounts were over-represented in the whole sample volume. The proposed work gave insights about the characterizations of these biased voters using Twitter data to analyze such political issues.

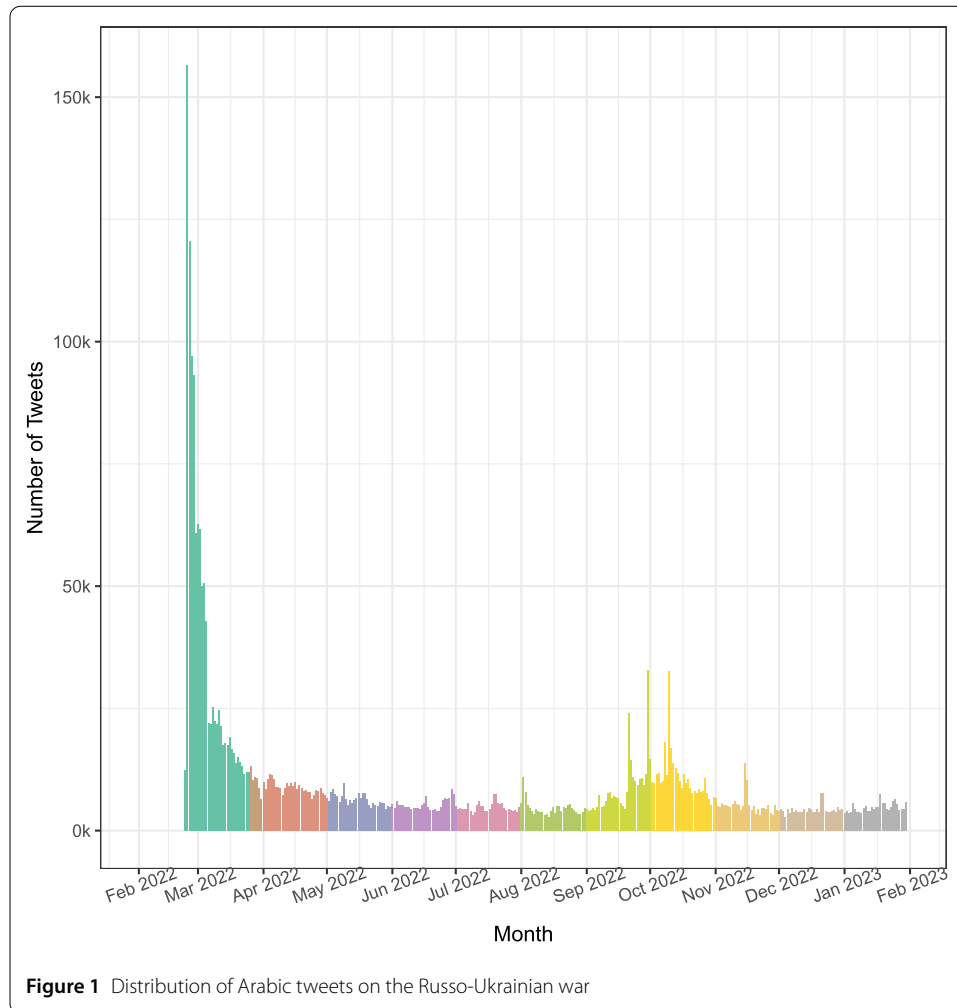
### 3 Methodology

In this section, we present our methodology in collecting tweets, data pre-processing, sentiment analysis, emotion analysis, and partiality analysis.

Generally, there are two main approaches for text analyses: (1) data-oriented approach based on machine learning methods and (2) more classical NLP approach based on lexicon analysis. Both paradigms go hand in hand, however, the current trend is leaned more towards the use of machine learning with the availability of more datasets and the huge successes of large language foundational models. However, Arabic is a low-resource language, where quality annotated large datasets are still missing. In addition, the computational resources needed to build or fine-tune pre-trained models are still huge and beyond the capabilities of many of the research institutions in the Arab world. So, we had to be careful in our decisions regarding the choice of the analysis paradigm. Sentiment analysis was the easiest; there is an abundance of work in the Arabic NLP literature that treats this problem and already there are several well-established pre-trained models on the Arabic language for sentiment analysis. So we resorted to the use of such models for supervised classification of the tweet's sentiment in a voting based criteria to reach a final decision regarding the target sentiment. Partiality analysis was the hardest. There are no pre-trained models, no annotated datasets, and no lexicon anywhere built for that purpose. So we even believe, to the best of our knowledge, that our work is the first methodical Arabic work handling this problem in any context. Therefore, we consider it our main contribution and we tried our best to take a data-oriented approach based on supervised classification. In order to do that, we did some tricks that include the following: (1) we filtered out large chunk of the tweets using the sentiment analysis part (removing neutral tweets), (2) we built an n-gram lexicon database, and (3) in addition to do some manual annotation for small part of the tweets using several subjects and taking their majority voting. All of these procedures at the end have been used to predict the bias of the major chunk of the tweets. The manual annotation is basically used for verification; and it was feasible as the number of classes were few, Pro-Russia, Pro-Ukraine, or neither. In addition, the decision on the class was rather easy, as it is not that subjective to determine the bias of the given tweet. So, as indicated, sentiment and partiality analyses were tightly coupled in our work. Emotion analysis was in the middle regarding its difficulty. Still there are no pre-trained or fine-tuned models for emotion analysis in the Arabic language. However, there is a constructed emotion lexicon that has already been used in published work. So, we used that lexicon for our analysis. We could not do the same procedures as in the partiality analysis as the number of classes here is too large for manual annotation (6 compared with 3 in partiality) In addition, it is much more subjective than the case of partiality making the annotation process more daunting and hungry for human resources.

#### 3.1 Collecting tweets

We have done tweets scraping starting from the 23rd of February 2022 (the day before the start of the war) till the 31st of January 2023. Every scrapped tweet contains the following raw information: Date-time, Tweet Id, Text (the text of the tweet including the tweet



hyperlink, any hashtags included in the tweet, and any mentioned account using the @ symbol in case of 'Replying to' other tweet). The language of the Tweets is of course Arabic. Figure 1 presents the number of Arabic tweets in our dataset concerning the Russo-Ukrainian War. It is clear from the figure that the rate of tweets kept decaying till nearly converging at a consistent level starting from June 2022 (except a relatively small spike in Sept - Oct 2022 "the beginning of Fall 2022" with the higher need to Oil worldwide and the inflation of its prices leading to more comments between Twitter users). This can be attributed to the initial unexpected turn of events at the beginning of the conflict, then a rather stability after the main motives, consequences, and outcomes have become clearer. In addition, there are durable periods in the conflict where the involved parties seem to be at stall.

The total number of tweets collected is 3,167,210 covering nearly the first 11 months of the still-ongoing conflict. As shown in Fig. 1, the peak in the number of tweets was on the 24th of February 2022 with over 156k tweets written in Arabic on that day; the day where Russia initiated the war. Again, as stated earlier another relatively smaller peak exists in Sept 2022 and Oct 2022, with the entrance of Fall 2022 and the higher demand for Oil worldwide and the inflation of its prices.

### 3.2 Data preprocessing

The main aim of the data preprocessing step is to present the text of tweets in a consistent form and reduce any potential noise (e.g., special symbols of hashtags). The data preprocessing procedure can be summarized in the following steps using ReGex in parsing and CaMeL-Tools to specifically deal with Arabic as follows:

- Removing usernames: any word starting with “@” is removed (e.g., @moe123).
- Removing links: any text starting with “www.” or “http” is removed.
- Removing emojis: any emoji in the tweet has been removed using the emoji’s unicode. (In the future, we will use these for further investigating emotions and bias.)
- Removing hashtags’ octothorpe, underscores, and hyphens (-) from tweets.
- Removing non-Arabic characters and words from tweets.
- Removing punctuation marks.
- Removing diacritics and elongations.
- Normalizing different forms of letters in Arabic to one consistent form; e.g., all the different forms of the Arabic letter Alef (أ, آ, إ) were converted to the unified form (ا).

Table 2 presents some examples of the data pre-processing procedure on a few tweets in our dataset with their English translation.

**Table 2** Data pre-processing procedure on a few tweets in our dataset

Before	After	English Translation
<p>Replying to @AA_S_AA1</p> <p>امريكا والغرب هم بغال للفرس والترك واسيادهم بني صهيون سلمو العراق لايران وقبيله الاحواز العربيه وفلسطين لليهود وقسمو باكستان وهم من يدعم كل عدو للعرب والمسلمين وخاصة نتاج الحزب الديمقراطي الابله بايدن وقبيله اليومه اوباما.</p> <p><a href="https://t.co/HMWzd7D3I5">https://t.co/HMWzd7D3I5</a></p>	<p>امريكا والغرب هم بغال للفرس والترك واسيادهم بني صهيون سلمو العراق لايران وقبيله الاحواز العربيه وفلسطين لليهود وقسمو باكستان وهم من يدعم كل عدو للعرب والمسلمين وخاصة نتاج الحزب الديمقراطي الابله بايدن وقبيله اليومه اوباما</p>	<p>America and the West are the mules of the Persians and the Turks, and their masters Zion's sons, gave Iraq to Iran, and before that, the Arab Ahwaz and Palestine to the Jews, and divided Pakistan and are the ones who support every enemy for the Arabs and Muslims, especially the Democratic party member, the idiot Biden, and before him the owl Obama</p>
<p>Replying to @andraoslb</p> <p>لا يسوغ موازنه ومقارنه بايدن المتهم بضعف قيادته اميركيا ودوليا ب بوتين الطاغيه الذي يجدد لنفسه ويسجن ويسمم أخصامه ومناقسيه السياسيين، ويغزو دولا مجاورة تكرر أ؛ ويهدد بمزيد من #الغزوات في أوروبا، ويهدد بالنووي الذي يقتل الملايين وقد ينهي البشرية! بوتين تجاوز في جنونه #هتلر</p>	<p>لا يسوغ موازنه ومقارنه بايدن المتهم بضعف قيادته اميركيا ودوليا ب بوتين الطاغيه الذي يجدد لنفسه ويسجن ويسمم أخصامه ومناقسيه السياسيين ويغزو دولا مجاوره تكرر أ؛ ويهدد بمزيد من الغزوات في اوروبا ويهدد بالنووي الذي يقتل الملايين وقد ينهي البشريه بوتين تجاوز في جنونه هتلر</p>	<p>It is not acceptable to equate and compare Biden, who is accused of his weak leadership in America and internationally by the tyrant Putin, who renews to himself, imprisons and poisons his opponents and politicians competitors, and invade neighboring countries repeatedly; and threatens by more invasions in Europe, and threatens by the nuclear weapon that kills millions and may end the humanity. Putin exceeds Hitler in his madness.</p>
<p>Replying to @US_World1</p> <p>في المقاطعة يجب من اذن ولي الامر في ذلك !. © هل اذن لهم #بايدن في مقاطعة الفودكا ؟ #الجاميه #سالم_الطويل <a href="https://t.co/469OYGCD4d">https://t.co/469OYGCD4d</a></p>	<p>في المقاطعة يجب من اذن ولي الامر في ذلك هل اذن لهم بايدن في مقاطعة الفودكا ؟ الجاميه سالم الطويل</p>	<p>In the boycott, it is a must to take the permission from the authority. Did Biden give them the permission to boycott the vodka? by Jamia Salem Al-Taweel</p>

### 3.3 Sentiment analysis

Identifying sentiment behind text is used to measure the attitude and feeling behind each tweet on the individual level, as well as to analyze the aggregate overall statistical pattern of the trending sentiments. In addition, such analysis can facilitate the subsequent other types of analyses. An example is identifying and consequently removing the 'Neutral'/'Mixed' tweets along with the 'Unspecified' ones (as these tweets are expected not to carry any emotions/biases towards the conflicting parties); then, utilizing the remaining attitudinal tweets (i.e., 'Positive' and 'Negative') in the partiality analysis step (i.e., 'Pro-Russia', 'Pro-Ukraine', or 'Neither').

In our 'Sentiment Analysis', the tweets are classified into four mutually exclusive labels: 'Positive', 'Negative', 'Neutral'/'Mixed', or 'Unspecified'. A tweet is labeled as 'Unspecified' when the three sentiment models, AraBERT, Mazajak, and CaMeL-Tools, annotate it with three different sentiments. It is interesting to note that 53.34% of the tweets were given the same label by the three sentiment models (assuming that 'Neutral' and 'Mixed' are the same). In other words, the three models exactly agreed on their sentiment decisions in more than half of the tweets.

### 3.4 Emotion analysis

Although 'sentiment' and 'emotion' are distinct notions that require distinct analyses, there is no agreed upon definition to distinguish between both [63]. Hence, our aim for emotion analysis is to determine the presence of words showing strong intolerance towards specific feeling(s). So, we take an operational stance towards such an analysis.

The main difference between 'sentiment' and 'emotion' analysis approaches used within this research is the 'emotion' dependence mainly on searching for previously specified n-grams within each tweet, while sentiment analysis is done using pre-trained models for identifying the sentiment. On one hand, these are two different technical paradigms (classical lexicon-based and data-oriented ML-based) to tackle two rather seemingly similar problems. On the other hand, it was not feasible to take a data-oriented approach to emotion analysis simply due to the lack of annotated quality emotion Arabic datasets and the lack of pre-trained models for such tasks in the Arabic language. However, we adopt a machine learning approach only for 'sentiment' analysis inspired by the abundance of work related to this aspect compared to 'emotion' analysis. Hence, the 'emotion' analysis task was done instead using a *lexicon-based* approach.

Each emotion has its own lexicon list<sup>1</sup> [64]. The lexicon is manually translated into Arabic from WordNet-Affect emotion lexicon [65], which is a subset of the English WordNet. Each entry in this lexicon is labeled with one of six emotions: Joy, Anger, Sadness, Fear, Surprise, and Disgust. The highest associated words with any of the aforementioned emotions is 'Joy' with 1156 words in Arabic followed by 'Anger' with 748 words as shown in Table 3.

Table 4 presents some examples of tokens for each emotion. This table includes the emotion itself (Anger, Disgust, Fear, Joy, Sadness, and Surprise), example tokens in Arabic, and example tokens in English.

We observed that the word 'war' (الحرب) is being repeated in thousands of tweets without showing any specific emotion behind. This is mainly caused by the fact that many tweets talking about the war, regardless of their opinion or how they feel, can be considered

---

<sup>1</sup><https://github.com/motazsaad/emotion-lexicon>.

**Table 3** Number of Tokens in Arabic for each emotion sorted in descending order

Emotion	Number of Tokens in Arabic
Joy	1156
Anger	748
Sadness	522
Fear	425
Surprise	201
Disgust	155
<b>Total</b>	<b>3207</b>

**Table 4** Example of tokens associated with each emotion

Emotion	Example Tokens in Arabic	Example Tokens in English
Anger الغضب	غيظ، حنق، سخط، غضب، حفيظة	wrath, umbrage, offense, pique, irritation
Disgust الاشمئزاز	اشمئزاز، مقت، نفور، تنافر، تفرز	repugnance, repulsion, revulsion, horror, nausea
Fear الخوف	شك، مريب، مشكوك، ريبة، خوف	unassertiveness, trepidation, timidity, timidness, timorousness
Joy الفرح	عبادة، توقير، تمجيد، تبجيل، إنتصار	victory, adoration, warmheartedness, warmth, triumph
Sadness الحزن	حزن، أسى، ويل، الدموع، أسف	woe, woefulness, weepiness, tear, regret
Surprise مفاجئة	عجب، تعجب، متعجب، مندهش، مذهش	wonder, wonderment, admiration, surprise, stupefactio

as stating news in a purely neutral way. Hence, we considered replacing it in the ‘Anger’ lexicon with ‘World War 3’ (الحرب العالمية الثالثة) and ‘The Great War’ (الحرب الكبرى).

Moreover, any tweet that was labeled with an emotion must contain at least two  $n$ -grams (max value of  $n$  is 3) from the specific emotion it was labeled with. If a tweet failed to meet this condition, either fewer than two  $n$ -grams per emotion or multiple emotions each with more than two  $n$ -grams, it would be labeled as ‘Null’ or ‘Mixed feelings’, respectively. The ‘Null’ class represents not having enough lexical strength for expressing any target emotion in the given tweet; whereas the ‘Mixed feelings’ class represents having enough richer lexical strengths for multiple feelings in the same tweet.

### 3.5 Partiality analysis

Partiality analysis means that the author of the tweet has some bias towards one of the two parties of the conflict. The most challenging step in ‘Partiality Analysis’ is to determine the amount of empathy each party receives across the data and the need for validation. There is rather lack of ‘Partiality Analysis’ pre-trained models in addition to the lack of annotated data. Pre-trained models may detect ‘Sentiment’, ‘Emotion’, or ‘Sarcasm’, but it will not necessarily be able to detect such partiality/bias.

Table 5 presents our methodology for bias analysis. Step 1 in Table 5 aims at building our own weighted  $n$ -grams lexicon.<sup>2</sup> A total of 223 uni-grams, bi-grams, and tri-grams have been collected in this lexicon with 118 considered ‘Pro-Ukraine’ ( $n$ -grams with positive weights) and 105 considered ‘Pro-Russia’ ( $n$ -grams with negative weights). The magnitude of the weight indicates the strength of the bias in either of the two directions. Table 6 presents some examples of particular  $n$ -grams and their weights.

<sup>2</sup>The weighted  $n$ -grams lexicon used in partiality analysis is currently available upon request.



**Table 5** Partiality analysis procedure

Step No.	Step
1.	Building a weighted lexicon of $n$ -grams directly related to the conflict.
2.	Filtering out the tweets with 'Neutral' and 'Unspecified' sentiment (their count is 1.379M tweets till Jan 31st 2023), as these tweets are expected not to carry any bias towards the conflicting parties. This first mandates collecting the Arabic tweets that are potentially linked to the conflict and classifying them according to their sentiment.
3.	Lemmatizing the remaining 1.789M tweets, then extracting those tweets containing the $n$ -grams, resulting in 449K tweets that will be used to train the LSTM neural network model in Step 5.
4.	Annotating each tweet of the 449K according to the pre-assigned weights as being 'Pro-Russia' (157K tweets), 'Pro-Ukraine' (140K tweets), or 'Neither' (152K tweets) ('Neither' means that the tweet could be irrelevant to the topic, showing strong opposition to both parties or contains relatively unsubstantial support to either party).
5.	Using the formerly identified tweets as a training dataset for a LSTM neural network model. This machine learning model will then annotate a testing dataset of about 1.471M tweets as being 'Pro-Russia', 'Pro-Ukraine', or 'Neither'.

**Table 6** Examples of particular  $n$ -grams and their weights

Arabic $n$ -gram	English Translation	Mean Score	Arabic $n$ -gram	English Translation	Mean Score
هجوم روسيا	Russia Attack	3	تقدم روسيا	Russia Progress	-3.75
حصار كييف	Kiev Siege	5.5	تحرير دونباس	Donbas Liberation	-5
القصف الروسي	Russian Bombing	5.75	انتصار روسيا	Russia Victory	-8.75
غزو روسيا	Russia Invasion	7.25	نفاق أمريكا	America Hypocrisy	-8.5
إرهاب روسيا	Russia Terror	9.75	نجاح روسيا	Russia Success	-7.25
الإرهاب الروسي	Russian Terrorism	9.75	الله يلعن الغرب	God damn the West	-8.25
استبعاد روسيا	Russia Exclusion	4.25	ادعاء أمريكا	America Claim	-7.5
جرائم روسيا	Russia Crimes	9.5	سيطرة روسيا	Russia Control	-4.75
جرم روسيا	Russia Offense	9.5	حرب استنزاف روسيا	Russia Attrition War	-3.5
جريمة روسيا	Russia Crime	9.5	مواجهة الغرب	Facing West	-4.25
استبعاد روسيا	Russia Exclusion	4.25	مواجهة أمريكا	Facing America	-4.25
سوريا	Syria	4	فيتنام	Vietnam	-2.5
معاقبة روسيا	Russia Punishment	5.25	عبيد الغرب	West Slaves	-8.75
سقوط روسيا	Russia Fall	5.75	عبيد أمريكا	America Slaves	-8.75
ديكتاتور روسيا	Russia Dictator	9	عنصرية الغرب	Western Racism	-8.25
ديكتاتور روسيا	Russia Dictator	9	عنصرية أمريكا	America Racism	-8.25
حظر روسيا	Russia Ban	5.75	شعارات أمريكا	America Slogans	-7
خسائر روسيا	Russia Losses	4.25	شعارات الغرب	Western Slogans	-7
الفشل الروسي	Russian Failure	7.75	نفاق الغرب	Western Hypocrisy	-8.5
الاحتلال الروسي	Russian Occupation	9	إرهاب الغرب	Western Terrorism	-9.5

These  $n$ -grams were collected by tracking press news, trending expressions used on social media, and the names of entities involved in the conflict. Each  $n$ -gram in the lexicon has a weight according to its intensity. For example, 'Putin is a hero' بولن بطل is considered a strongly 'Pro-Russia' bi-gram. On the other hand, 'Russia Terrorism' إرهاب روسيا is considered a strongly 'Pro-Ukraine' ('Anti-Russia') bi-gram. The weights range in the interval from -10 (highest Pro-Russia) to +10 (highest Pro-Ukraine) and are the result of a majority voting (mean value) of four different individuals each independently assigning a weight for every  $n$ -gram in this lexicon. These four individuals are chosen randomly and all have neutral stance towards the conflict. Using this weighted  $n$ -grams lexicon, we identify and annotate tweets that are directly related to the dispute. These tweets were utilized afterwards as a training dataset for a machine learning model (LSTM neural network). This ML model afterwards annotates a testing tweets dataset. The LSTM machine learning model has the ability to extract features from the annotated training dataset without

**Table 7** Partiality analysis tweets of each kind ('Pro-Russia', 'Pro-Ukraine', 'Neither')

Example in Arabic	English Translation	Score	Label
جريمة روسيا في اوكرانيا لم تحدث في التاريخ	The Russian crime in Ukraine has never happened in history	9.5	'Pro-Ukraine'
هدف روسيا هو تحرير دونباس من ايدي الاوكران واعوانهم	The Russian aim is to liberate Donbas from Ukrainian hands and their allies	-5	'Pro-Russia'
القصف الروسي وحصار كييف مستمر والعالم كله مع اوكرانيا كان فين ادعاء امريكا ده وقت ما اولادنا كانوا بيتقتلوا الله يلعن الغرب	Russian bombing and Kiev's siege continues, and all the world stands with Ukraine. Where was this America's claim in the period where our children were being killed? God damn the west.	-4.5	'Pro-Russia'
الوضع الحالي في كييف سي للغاية	The Current State in Kiev is so bad.	0	'Neither'

using hand crafted features (e.g., partiality  $n$ -grams lexicon) and thus is able to learn inherently the context of the tweets. In addition, the rather small training dataset (that contains the  $n$ -grams partiality lexicon) is constructed beforehand instead of annotating the whole dataset (using only the lexicon-based approach).

Step 2 in Table 5 aims at excluding the 'Neutral' and 'Unspecified' tweets. This is mainly attributed to removing noise and irrelevant tweets (e.g., news from outlets, spams, etc.). Recall from the 'Sentiment Analysis' procedure that 'Neutral' tweets are determined from the majority voting of the three sentiment models as 'Neutral', while 'Unspecified' is assigned when the three sentiment models annotate the tweet with three different sentiments.

Step 3 in Table 5 aims at mitigating the various possible word forms in Arabic and overcoming the Arabic's rich morphology. Thus, the lexicon illustrated in Step 1 and the dataset obtained from Step 2 (about 1.789M tweets) were first lemmatized resulting in about 449K tweets. Lemmatization was done using the Farasa lemmatizer [14] which is a fast Arabic segmenter based on SVM-rank with linear kernels.

Step 4 in Table 5 is further illustrated in Table 7 that presents samples of tweets in Arabic, their English translations, and the bias of each tweet. Notice the magnitude of the score indicates the strength of the bias towards the corresponding party. Thus, we show not only the class, but the strength of that class.

Notice in Table 7, the third example contains  $n$ -grams favoring Ukraine and others favoring Russia. However, the whole tweet was labeled 'Pro-Russia'. This is mainly because the  $n$ -grams favoring Russia, e.g., 'God damn the west' الله يلعن الغرب have higher weights than the  $n$ -grams favoring Ukraine as 'Russian bombing' القصف الروسي and 'Kiev's Siege' حصار كييف; thus the tweet was annotated as 'Pro-Russia'. We assume that the  $n$ -gram(s) presence in a tweet implies that the author of this tweet favors one side of the conflict. Accordingly, each tweet in the dataset was given a final weight based on the summation of the  $n$ -grams weights present in the tweet. Specifically, if the tweet's total weight is positive, it is annotated as 'Pro-Ukraine', while if the summation is negative, it is annotated as 'Pro-Russia', else if the summation is zero, it is annotated as 'Neither'. The magnitude of the score indicates the strength of the bias towards the corresponding party.

To validate our results, in particular, the effectiveness of our construction of the partiality lexicon, 400 tweets were randomly selected and manually annotated with the proper partiality orientation, in order to compare with the results of the  $n$ -grams lexicon-based approach. These 400 tweets were assigned to four individuals for manual labeling as 'Pro-

Russia' or 'Pro-Ukraine', then a majority voting was taken for each tweet based on the labeling of the four individuals. It turned out that 81.25% of the tweets labels were consistent with the labels assigned by the  $n$ -grams lexicon-based approach.

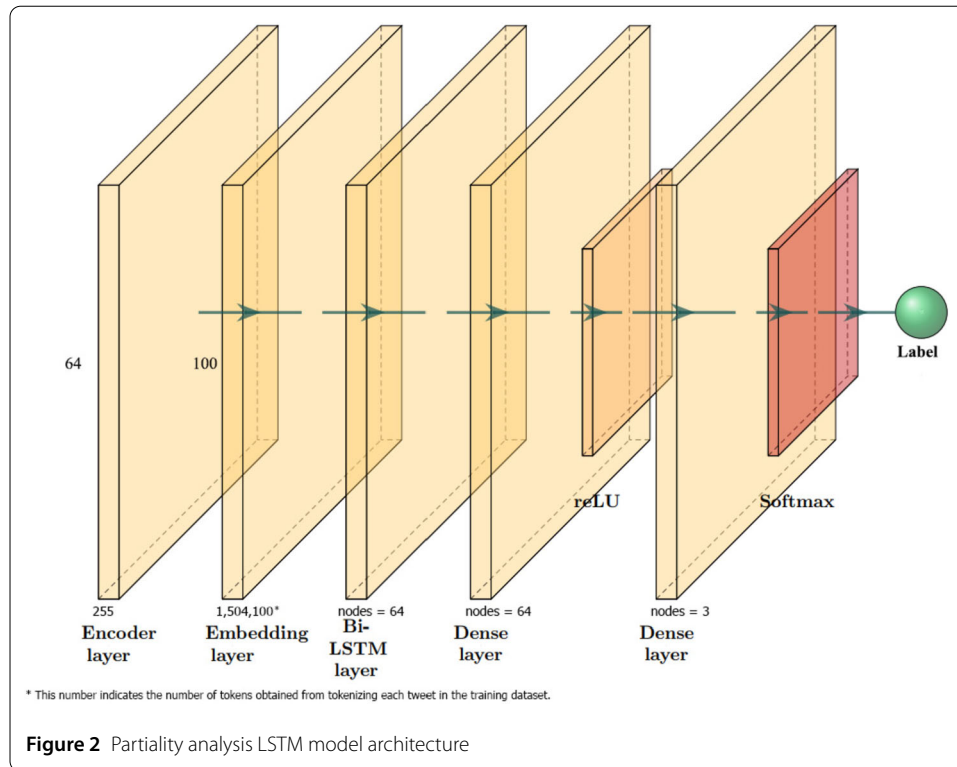
The disagreements in the manual annotation and the lexicon-based annotation are noticed strongly within tweets containing countries or personnel involved in the conflict and can exist in one context criticizing Russia or another context criticizing Ukraine. As an example, countries like Syria سوريا may be used in a biased context (e.g., 'Putin destroyed Syria' سوريا دمر بوتين) or a less-biased/neutral context (e.g., 'Russia made Ukraine a new Syria' روسيا جعلت من اوكرانيا سوريا جديدة). The first tweet indicates vigorous expression 'destroyed' دمر making the tweet classified easily as 'Pro-Ukraine' by both the manual and lexicon-based annotations. For the second tweet, the criticism of Russia is less clear with no vigorous expression like the first sentence (annotated as 'Neither' by the lexicon-based annotation). However, the terrible status of Syria since the Russian intervention gives a strong indication that the belief behind this sentence is 'Pro-Ukraine'. Considering personnel involved in the conflict, some can be considered as 'Pro-Ukraine' (e.g., 'Putin is a criminal' بوتين مجرم), however, in another context it is 'Pro-Russia' (e.g., 'Some people say Putin is criminal while he is not' البعض يقول بوتين مجرم وهو ليس كذلك). Thus, we depend on the LSTM machine learning model as a more 'data science' approach (i.e., learning from the  $n$ -grams and the tweet context itself). Removing such tweets containing countries or personals that can be used in a context criticizing Russia or another context criticizing Ukraine (e.g., 'Syria' سوريا, 'Afghanistan' افغانستان, 'Vietnam' فيتنام, 'Putin' بوتين, 'Zielinski' زيلينسكي) lead to an accuracy increase from 81.25% to 85.43%. The full list of countries and personals involved in the conflict that were removed leading to such accuracy increase is presented in Table 8.

In Step 5 in Table 5, after collecting and searching about 449K tweets containing the  $n$ -grams that indicate favoring either party, these tweets were used to train an LSTM model in order to classify a testing dataset of about 1.471M tweets as being either: 'Pro-Russia', 'Pro-Ukraine', 'Neither'. The architecture of this LSTM model is shown in Fig. 2.

The 'Text Vectorization' layer represents how a raw text is encoded into a numerical form to be given to the model representing each individual word in the tweet as a vector of fixed length. After doing a comprehensive review within the literature we chose a particular representation of the sentences that is inspired by the work in [11] which requires a reasonable amount of computational resources as compared to that required by the BERT model [14, 16]. Sentences are represented as a two dimensional embedding matrix where

**Table 8** Conflicting  $n$ -grams causing disagreements in the manual annotation and lexicon-based annotation

Conflicting $n$ -gram	English Translation
بوتين	Putin
زيلينسكي	Zielinski
امريكا	U.S
افغانستان	Afghanistan
فيتنام	Vietnam
الصين	China
فلسطين	Palestine
العراق	Iraq
سوريا	Syria

**Table 9** Partiality analysis LSTM hyperparameters

Parameter	Value
Epochs	5
Optimizer	Adam
Learning Rate	0.001
Loss Function	Categorical Crossentropy

each row represents a word, and each word is represented by the corresponding word embedding.

The Embedding matrix mentioned in the architecture is critical in order to ensure the Arabic tweets are properly encoded while being fed to the LSTM layer. It is based on the skip-gram using 100M tweets used in Mazajak [11]. Bidirectional LSTM layer allows the model to capture the long-term dependencies and context of the text. We have not used CNN layers for feature extraction as the embedding matrices already contain enough feature content, so instead we connected the LSTM layers directly to a dense layer with ReLU activation function. Our model showed 95.07% test accuracy and 95.21% validation accuracy after running it with the hyperparameters shown in Table 9.

Beside training an LSTM model from scratch, we fine-tuned an AraBERT model on the same task of labeling a tweet as either 'Pro-Russia', 'Pro-Ukraine', or 'Neither' on the same dataset. The hyperparameters of fine-tuning are presented in Table 10. The difference between the hyperparameters used in training the LSTM model and the ones used in fine-tuning the AraBERT model was the lesser number of epochs and the lower learning rate. This is mainly attributed to the recommendation of the BERT authors. Referring to [66], in order to achieve good performance across all tasks, the number of epochs can be 2, 3, or 4, and the learning rate values for the Adam optimizer can be  $5e-5$ ,  $3e-5$ , or  $2e-5$ .

**Table 10** Hyperparameters used in fine-tuning AraBERT for the ‘Partiality Analysis’ task

Parameter	Value
Epochs	3
Optimizer	Adam
Learning Rate	5e-5
Loss Function	Categorical Crossentropy

**Table 11** Sentiment labeling obtained with majority voting of three different pre-trained foundational models: (1) Mazajak [11], (2) AraBERT [14], and (3) the CaMeL-Tools [16]

Sentiment	No. of Tweets	Percentage
Negative	1,679,300	53.0%
Neutral	1,319,808	41.7%
Positive	109,333	3.45%
Unspecified	58,769	1.86%

The results of fine-tuning the AraBERT model showed that it achieved a test accuracy of 94.69% and a validation accuracy of 94.64%. In order to compare the results of both the LSTM model and the fine-tuned AraBERT model, 50,000 unlabeled tweets were chosen at random for labeling. Both models agreed on the annotation of 45,113 tweets which means that both models gave the same label for over 90% of the chosen sample.

## 4 Experimental work

### 4.1 Sentiment analysis

We have collected 3,167,208 tweets starting from Feb 23rd, 2022 till Jan 31st 2023. The results of the sentiment analysis are shown in Table 11.

It is obvious that ‘Negative’ and ‘Neutral’ labels dominate the attitude comprising nearly 95% of the tweets. Most of the tweets (53%) showed a ‘Negative’ attitude towards the war. This can be attributed to two things: (1) war by itself is a negative human experience involving casualties, specially for civilians and the destruction of human civil facilities; so civilians specially suffer the most from wars and (2) pragmatic reasons due to the direct negative effects of the war on the economic and lifestyle for people in most of the Arabic-speaking region. In addition a large percentage (about 40%) have a ‘Neutral’ attitude towards the conflict. These could be oblivious to the conflict as it is geographically happening at a distance and the relevant parties have no strong ties to the region. Some people could also be a little ignorant of the dire consequences of this conflict over the region, namely, at least, high inflation rates, the crisis in food supplies, and the soaring increase in energy prices. A small fraction of the people (3.45%) have a positive attitude towards this conflict. They could be originating from people in the Gulf area where they benefited much economically from the ongoing war. However, it is worth thoroughly studying why such an attitude is happening.

Table 12 presents examples of sentiment tweets of each kind (Positive, Negative, Neutral). Figures 3 and 4 illustrate the evolution of each individual sentiment over time since the beginning of the war and its normalized version, respectively. It is important to note that the ‘Unspecified’ labeled tweets were removed before plotting.

Table 13 presents some examples for the change in each sentiment over time, that is the difference between each two consecutive points. Figure 5 illustrates the evolution of this difference over time. It is clear from the table and the figure that at the beginning of

**Table 12** Sentiment example tweets of each kind ('Positive', 'Negative', 'Neutral')

Example in Arabic	Translation in English	Label
اهلا وسهلا بالضيف الثقيل راعي العلوم الطويله.. هذا الشخص محنك سياسيا، يعرف إقناع الأصدقاء، والخصوم تجده مبتسم في اشد حالات الغضب، لهذا التاريخ اليوم يسجل اسم #سيرجي_لافروف دورة بالأحداث العالمية، يسعى في إعادة روسيا الي الواجهة،حتى وإن تسبب ذلك بالنشوب #الحرب_العالمية_الثالثة	Welcome to the heavy guest, the patron of long sciences.. This person is politically savvy, knows how to persuade friends, and opponents find him smiling in the most angry states. For this date today, the name #Sergey_Lavrov records his role in world events, seeks to bring Russia back to the fore, even if this causes the outbreak of #World_War_Three	'Positive'
المؤامرة الي بعد كورونا هي تجويع الشعوب كما توقع عدة محللين وعدة فلكيين... #الحرب_العالمية_الثالثة #مؤامرة	The conspiracy after Corona is to starve the people, as predicted by several analysts and several astronomers... #World_War_Three #Conspiracy	'Negative'
نصف صادرات روسيا الكلية، من النفط تم منعها من دخول الاسواق الاوروبية ! يجري الروس مباحثات الآن مع الدول الاسيوية للبحث عن اسواق بديلة.. #روسيا #الحرب_العالمية_الثالثة	Half of Russia's total oil exports were prevented from entering European markets! The Russians are now in talks with Asian countries to search for alternative markets. #Russia #World_War_Three	'Neutral'

**Table 13** Differences of sentiments of each kind ('Negative', 'Neutral', 'Positive')

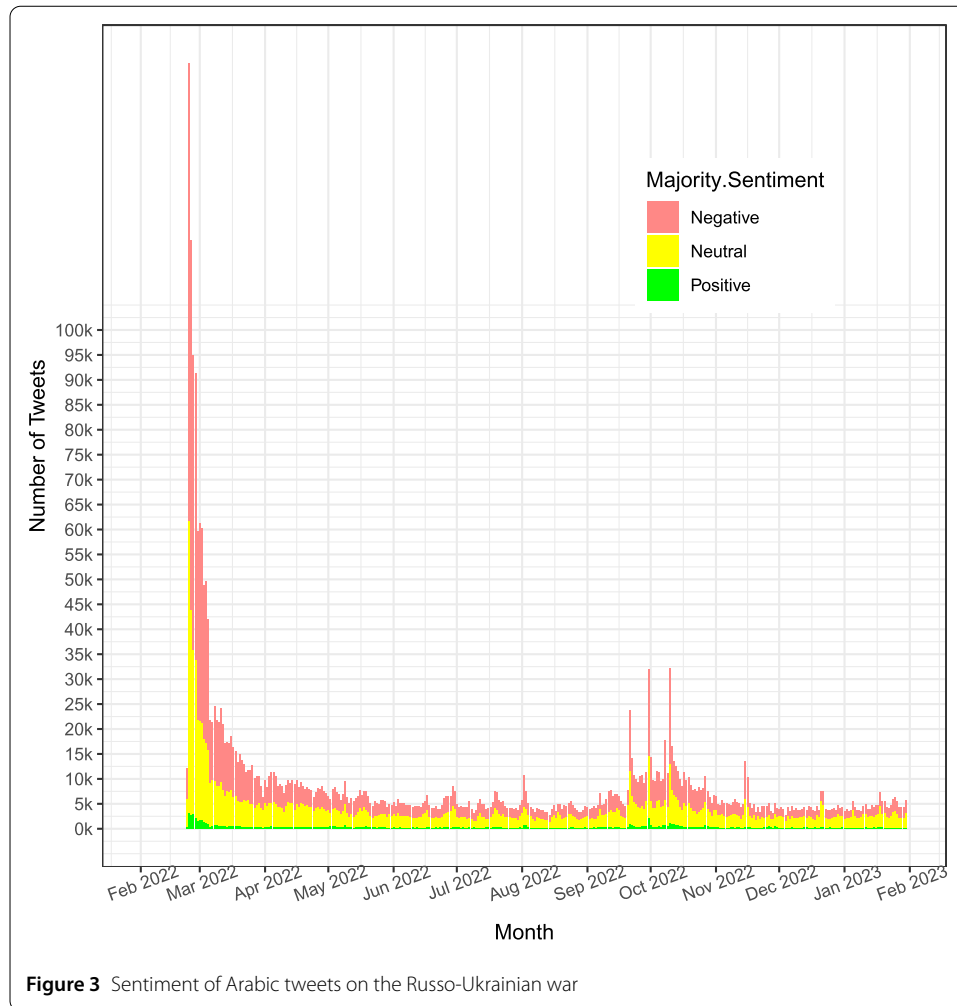
Sentiment	Date	Count	Derivative
Negative	2022-02-23	6308	NA
Negative	2022-02-24	91,889	85,581
Negative	2022-02-25	74,225	-17,664
Negative	2022-02-26	58,970	-15,255
Negative	2022-02-27	57,364	-1606
Negative	2022-02-28	37,730	-19,634
Neutral	2022-02-23	5596	NA
Neutral	2022-02-24	58,445	52,849
Neutral	2022-02-25	41,154	-17,291
Neutral	2022-02-26	33,065	-8089
Neutral	2022-02-27	31,671	-1394
Neutral	2022-02-28	20,239	-11,432
Positive	2022-02-23	267	NA
Positive	2022-02-24	3118	2851
Positive	2022-02-25	2650	-468
Positive	2022-02-26	2833	183
Positive	2022-02-27	2184	-649
Positive	2022-02-28	1595	-589

the war the Negative sentiments increased intensively then decreased dramatically in the first few days. This may be attributed to the initial shock at the beginning of the war that lightened after a few days and started oscillating between slight increase and decrease.

Notice that as the hype of the events decreased overtime, tweets relevant to the subject faded away from being the trending subjects (except relatively small spikes in Sept 2022 and Oct 2022, i.e., the entrance of Fall 2022 and the higher need for Oil worldwide and the inflation of its prices; manifested in the high increase in the 'Negative' sentiment). The shrinkage can be seen to follow a power law pattern ( $\propto \frac{1}{k^d}$ , for some positive  $d$ , and  $k$  represents the time, for example, in days) since the beginning of the conflict. The total number of relevant tweets as well as the corresponding sentiments stabilize nearly after two months since the beginning of the conflict.

The most impressive thing we notice in Fig. 3 is how big was the difference between the 'Negative' and 'Neutral' tweets in the beginning of the war and how it decreased till Jan 2023 (again except the beginning of Fall 2022 for the reasons stated earlier). In pro-

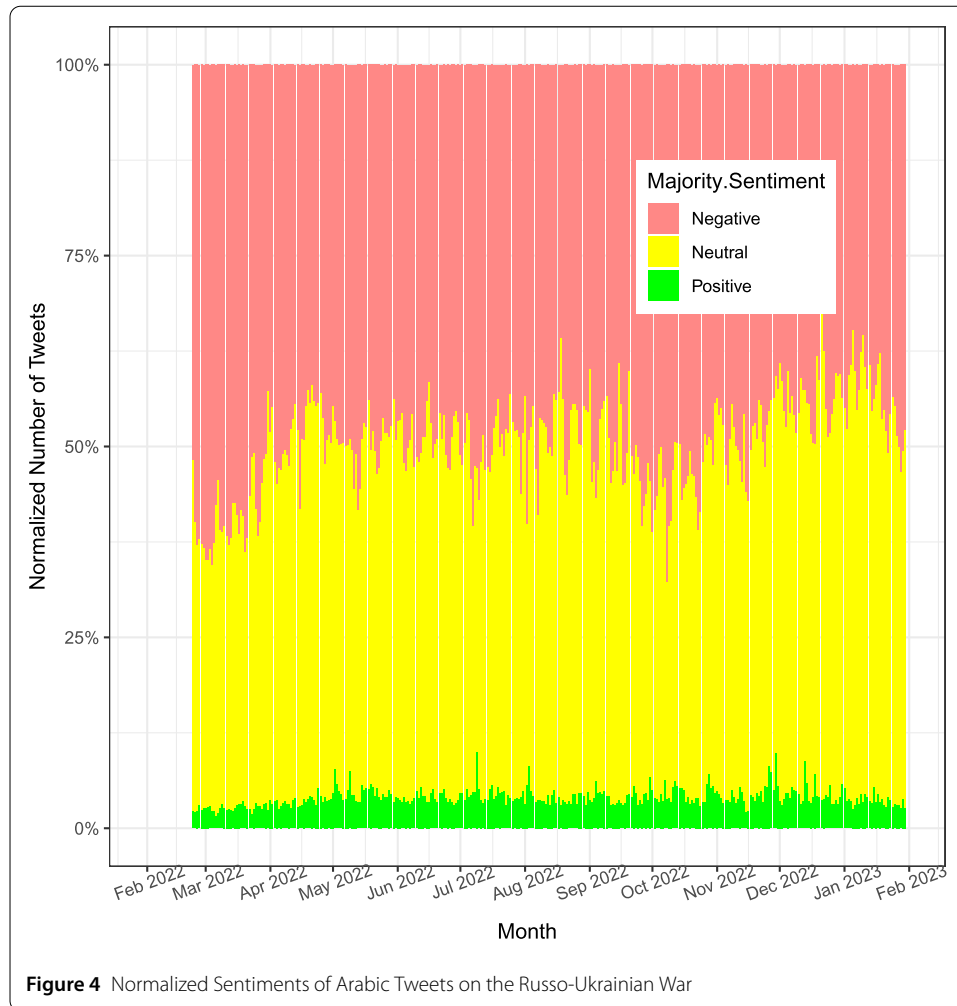




portion to the total number of tweets, the difference between the ‘Negative’ and ‘Neutral’ tweets was about 25% in February and March 2022, then decreases to below 10% in Jan 2023; specifically the percentage of ‘Negative’ has decreased from 62% to 44% while the percentage of ‘Neutral’ has increased from 36% to 53%.

#### 4.2 Emotion analysis

In emotion analysis, a ‘Null’ label indicates that the tweet contains less than two  $n$ -grams from the emotion lexicon. A ‘Mixed’ label indicates that the tweet contains enough  $n$ -grams from at least two different emotions (e.g., ‘Anger’ and ‘Sad’). 89.1% of the tweets did not contain any  $n$ -grams from the emotion lexicon, consequently these tweets were annotated as ‘Null’. This may be attributed to writing with different dialects or a rather limited size of the currently adopted lexicon. As shown in Table 3, the total number of tokens in Arabic is 3207 that is considered for future extension. The results of emotion analysis are shown in Table 14, which can be extrapolated to uncharted tweets due to limited lexicon size. It is apparent that there is no strong feeling as the conflict may seem a bit far from next door. Also, there are some many tweets with mixed feelings indicating the perplexity towards that conflict and the confusion it causes amongst the public; or due to naturally co-occurring emotions such as fear and anger. Generally, we can say that

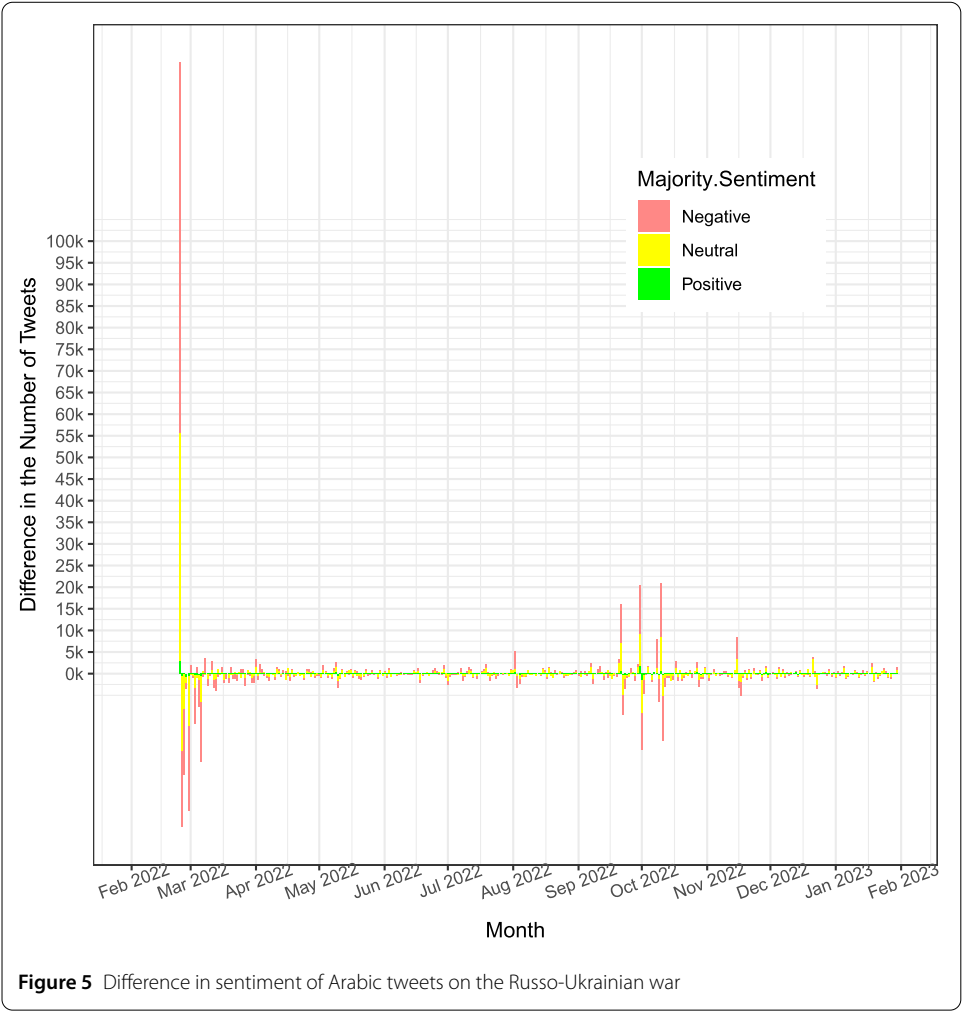


nearly 10.93% of the tweets were considered to contain strong expressions of emotion (i.e., Anger, Joy, Fear, Sad, Surprise, Disgust, or Mixed Feelings).

Table 15 dives deeper into the ‘Mixed Feelings’ category, showing the frequency of each emotion in the tweets labeled as ‘Mixed Feelings’. Each row gives the number of tweets in the mixed feelings category that carry the given emotion listed in the first column. For example, among the 234,384 of ‘Mixed feelings’ tweets, 174,548 of them carry an ‘Anger’ feeling. It is clear from the table that ‘Anger’, ‘Sad’, and ‘Fear’ are the most frequent emotions in the ‘Mixed feelings’ category. These emotions are no doubt the most relevant emotions with any violent conflicts such as wars.

Table 16 presents the results of different combinations of all emotions. In this table the combination (‘Anger’, ‘Sad’) is the most frequently occurring combination followed by (‘Anger’, ‘Fear’). These co-occurring emotions are natural as, for example, anger is usually accompanied with any of the emotions of sadness and/or fear; similarly, for the other combinations. This validates and indicates the efficacy of both our tweets emotion annotation scheme and the effectiveness of the developed classification models.

For a deeper dive in a sample emotion (like ‘Disgust’) in order to study its presence with all other emotions; interestingly, ‘Anger’ then ‘Fear’ are the most present emotions with ‘Disgust’ that can be attributed to a person feel angry or fear while expressing a disgusting



**Table 14** Emotion analysis obtained according to the presence of different *n*-grams in each tweet

Emotion	No. of Tweets	Percentage
Null	2,820,959	89.100%
Mixed Feelings	234,384	7.400%
Joy	53,595	1.690%
Anger	24,400	0.770%
Fear	23,538	0.743%
Sad	5673	0.179%
Surprise	2353	0.074%
Disgust	2308	0.073%

attitude. For instance, the couple ‘Disgust’ and ‘Anger’ comes in 21,560 tweets from all the 234,384 tweets annotated as ‘Mixed feelings’ (about 9.19%). While, the couple ‘Disgust’ and ‘Fear’ comes in 7138 tweets (about 3.04%); notice that the same tweet that is annotated as ‘Mixed feelings’ can have the triple ‘Disgust’, ‘Anger’, and ‘Fear’.

Table 17 presents sample tweets with the emotion of each kind: ‘Anger’, ‘Joy’, ‘Fear’, ‘Sad’, ‘Surprise’, and ‘Disgust’. The table includes an example in Arabic, the translation in English, and the emotion annotation. As stated earlier in Sect. 3.4; each emotion has its own tokens list; for some examples, the reader can refer to Table 4.

**Table 15** 'Mixed feelings' in emotion analysis obtained according to the presence of different *n*-grams in each tweet. The size of 'Mixed feelings' is 234,384

Emotion	No. of Tweets	Percentage
Anger	174,548	74.47%
Sad	110,329	47.07%
Fear	105,113	44.84%
Joy	87,140	37.17%
Disgust	32,793	13.99%
Surprise	25,479	10.87%

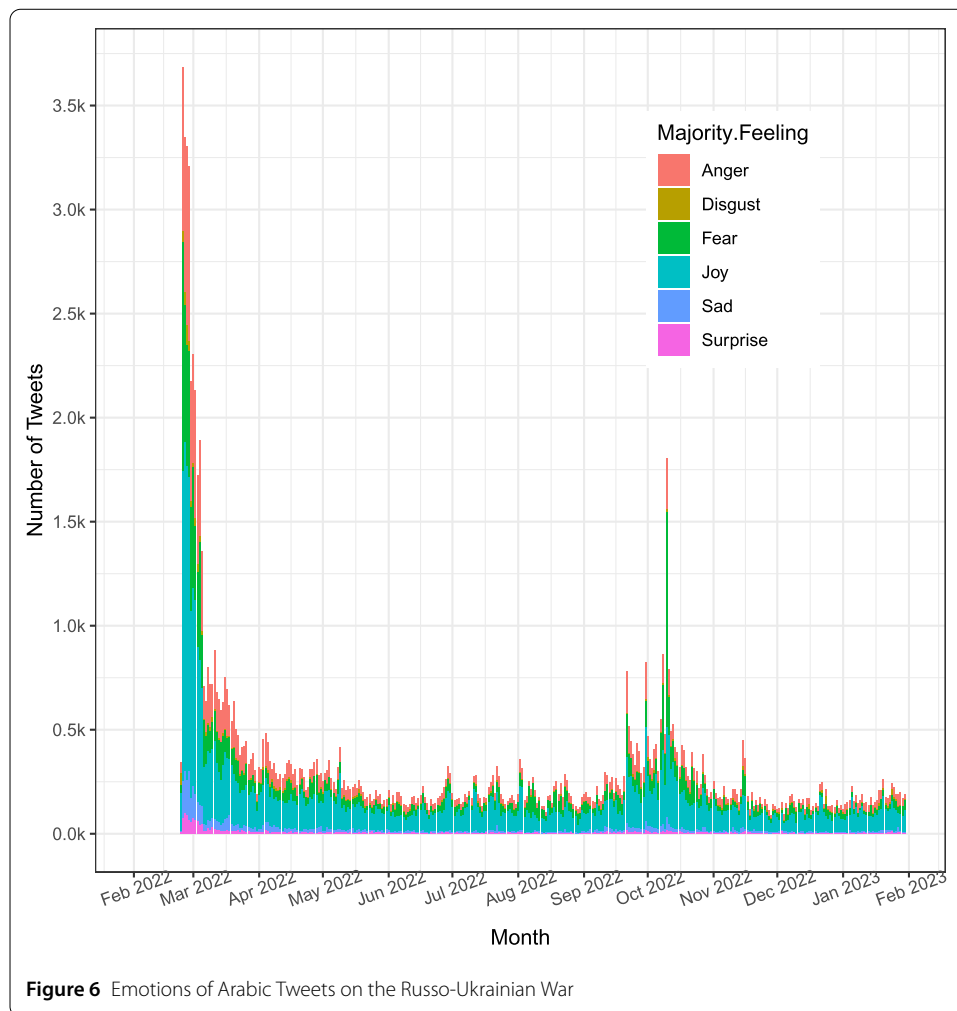
**Table 16** Different combinations of all emotions obtained according to the presence of different *n*-grams in each tweet where the total size of 'Mixed feelings' is 234,384

Emotion 1	Emotion 2					
	Joy	Disgust	Anger	Fear	Sad	Surprise
Joy	-	11,327 (4.83%)	46,205 (19.70%)	32,844 (14.01%)	23,458 (10.01%)	11,097 (4.73%)
Disgust	11,327 (4.83%)	-	21,560 (9.19%)	71,38 (3.04%)	6393 (2.72%)	2699 (1.15%)
Anger	46,205 (19.7%)	21,560 (9.19%)	-	68,053 (29.03%)	88,679 (37.8%)	9526 (4.06%)
Fear	32,844 (14.01%)	7138 (3.04%)	68,053 (29.03%)	-	33,240 (14.18%)	8964 (3.82%)
Sad	23,458 (10.01%)	6393 (2.72%)	88,679 (37.8%)	33,240 (14.18%)	-	4995 (2.13%)
Surprise	11,097 (4.73%)	2699 (1.15%)	9526 (4.06%)	8964 (3.82%)	4995 (2.13%)	-

**Table 17** Sample tweets with emotion of each kind ('Anger', 'Joy', 'Fear', 'Sad', 'Surprise', 'Disgust')

Example in Arabic	Translation in English	Label
ما يحدث في العالم الان ضرب من الجنون. جنون الاوبئة واللقاحات. جنون ارتفاع الاسعار. جنون الحروب. جنون رهف القحطاني. وكأنك تعيش احلام اليقظة #الحرب_العالمية_الثالثة	What is happening in the world right now is crazy. Epidemics and vaccines madness. Crazy high prices. Wars' madness. Rahaf Al-Qahtani Madness. It is like you are daydreaming. #World_War_Three	'Anger'
لا أحد ينتصر في الحرب سوى الحرب نفسها، فهي التي تنتصر على الجميع. #الحرب_العالمية_الثالثة #عينك_على_العالم	No one wins war but war. The war itself is the one that wins all. #World_War_Three #Your_eye_is_on_the_world	'Joy'
تهديد التلفزيون الروسي الحكومي بإغراق بريطانيا في موجة مد عاتية (تسونامي) متعمدة عبر انفجار نووي تحت البحر قبالة السواحل البريطانية. #روسيا #الحرب_العالمية_الثالثة #صاروخ_بريطانيا	The threat of Russian state television to plunge Britain into a deliberate tidal wave (tsunami) through an undersea nuclear explosion off the British coast. #Russia #War #World_War_Three #Missile #Britain	'Fear'
جيش هزيل للأسف لم تحرروا لا خاركييف ولا اوديسا ولا نيكولايف منذ شهور	A meager army. Unfortunately, you have not liberated Kharkiv, Odessa, or Nikolaev for months	'Sad'
بوتين بطل كشف زيف الاعلام كشف زيف الرياضة كشف العنصرية كشف النازية كشف خوات القحبة خليه بحرقهم	Putin is a hero who exposed the falsehood of the media, exposed the falsehood of sports, exposed racism, exposed Nazism, exposed the sisters of whores, let him burn them.	'Surprise'
تمرد في جيش بوتين الخاص: نرفض القتال في أوكرانيا #روسيا #أوكرانيا	Rebellion in Putin's private army: "We refuse to fight in Ukraine" #Russia #Ukraine	'Disgust'

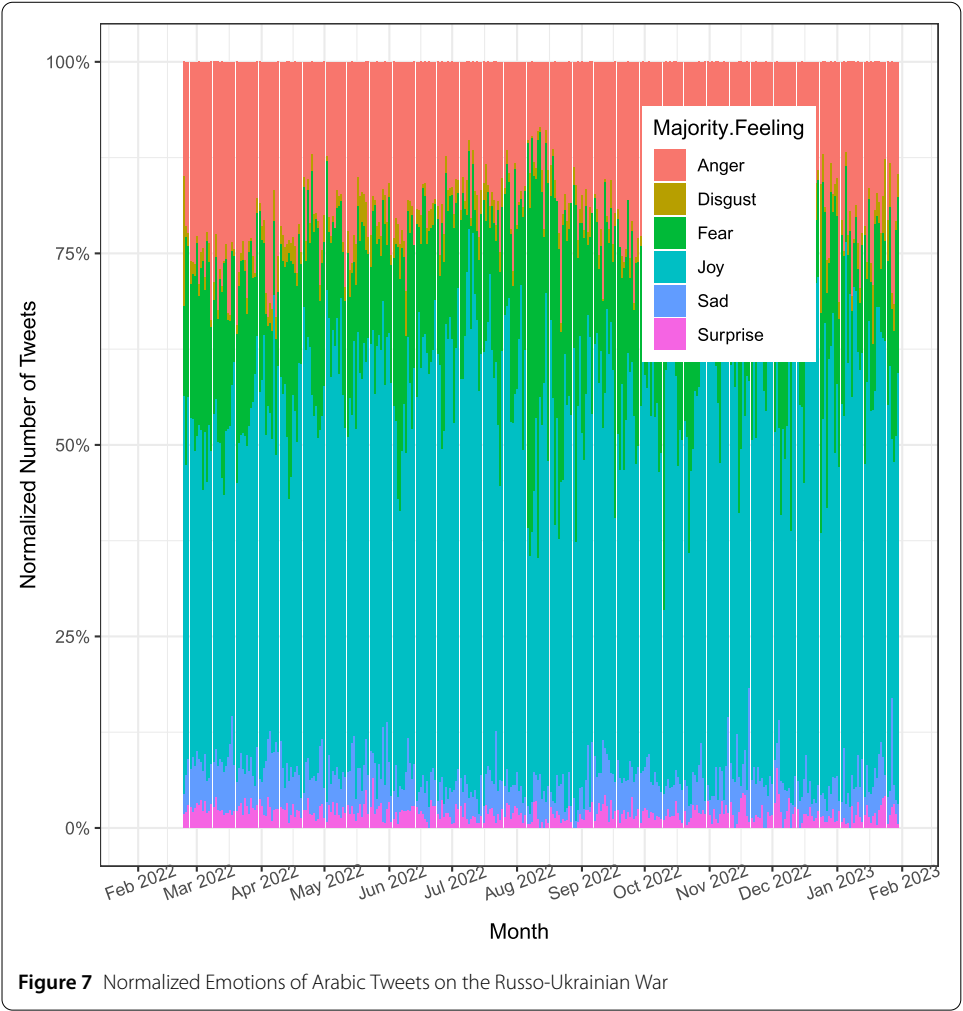
Figures 6 and 7 illustrate the evolution of each individual emotion over time since the beginning of the war and its normalized version, respectively. Similar to the tweets temporal trend shown in Fig. 1, the amount of emotions kept decaying till nearly converging at a consistent level starting from June 2022. Again, this can be attributed to the initial unexpected surprise at the beginning of the conflict, then a rather stability after the



main motives, consequences, and outcomes become clearer. However, starting from late September there is another upsurge in the emotions. This may be attributed to the new tide of the war in the opposite direction with the counter-attack of Ukraine seizing territories from Russian troops that suffered several defeats and started receding back. Within this decaying pattern, still the emotions' majority is either 'Joy' or 'Anger'. This is mainly attributed to the highest number of Arabic tokens in the adopted emotion lexicon being associated with 'Joy' (1156 tokens) followed by 'Anger' (748 tokens) as presented previously in Table 3. So, there is some sort of bias leading to the emotion lexicon needing to be enhanced which is put in our future work. In addition, the 'Joy' emotion was found to be existing in tweets related to the war but including hashtags relevant to other emotional subjects in the Arab region, e.g., *من قلبي للقدس سلام* (Peace From my heart to Jerusalem) and *الاتحاد النصر* (Union Victory). This can be attributed to increase the tweet visibility by mentioning more than one hashtag (especially trending ones) in the tweet (even if not being relevant to the Russo-Ukrainian war).

#### 4.3 Partiality analysis

As mentioned in Table 5, about 1.379M tweets with 'Neutral' or 'Unspecified' sentiments were initially filtered out. The remaining 1.789M tweets include about 449K classified by



**Figure 7** Normalized Emotions of Arabic Tweets on the Russo-Ukrainian War

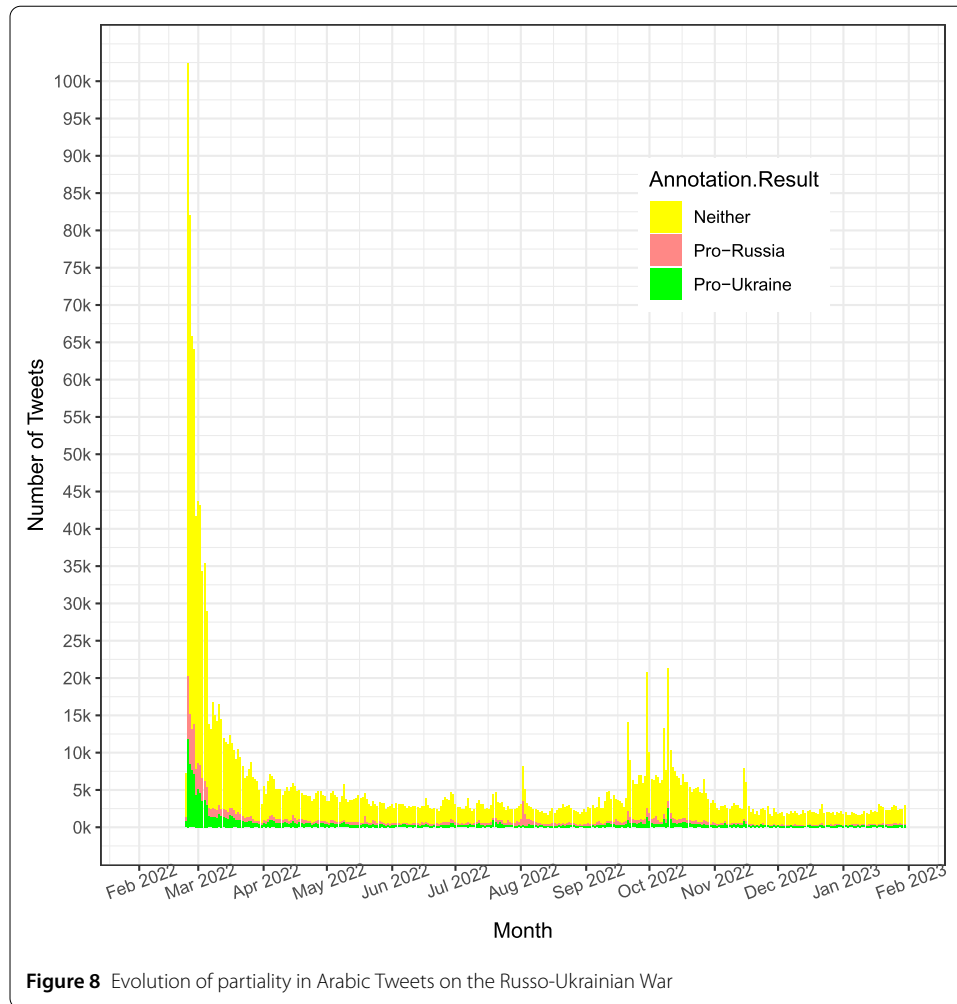
**Table 18** Partiality analysis obtained according to Step 4 and Step 5 mentioned in Table 5

Partiality	Lexicon training set	LSTM testing set	Total no. of Tweets	Percentage
Pro-Russia	157K	13K	170K	8.85%
Pro-Ukraine	140K	50K	190K	9.90%
Neither	152K	1.408M	1.560M	81.25%

the lexicon-based approach as following: ‘Pro-Russia’ (157K tweets), ‘Pro-Ukraine’ (140K tweets), or ‘Neither’ (152K tweets). Using the formerly identified tweets as a training dataset for an LSTM neural network model; the results of the testing dataset; ‘Pro-Russia’ (13K tweets), ‘Pro-Ukraine’ (50K tweets), and ‘Neither’ (1.408M tweets). The final results of the ‘partiality analysis’ are shown in Table 18; ‘Pro-Russia’ (170K tweets), ‘Pro-Ukraine’ (190K tweets), or ‘Neither’ (1.560M tweets).

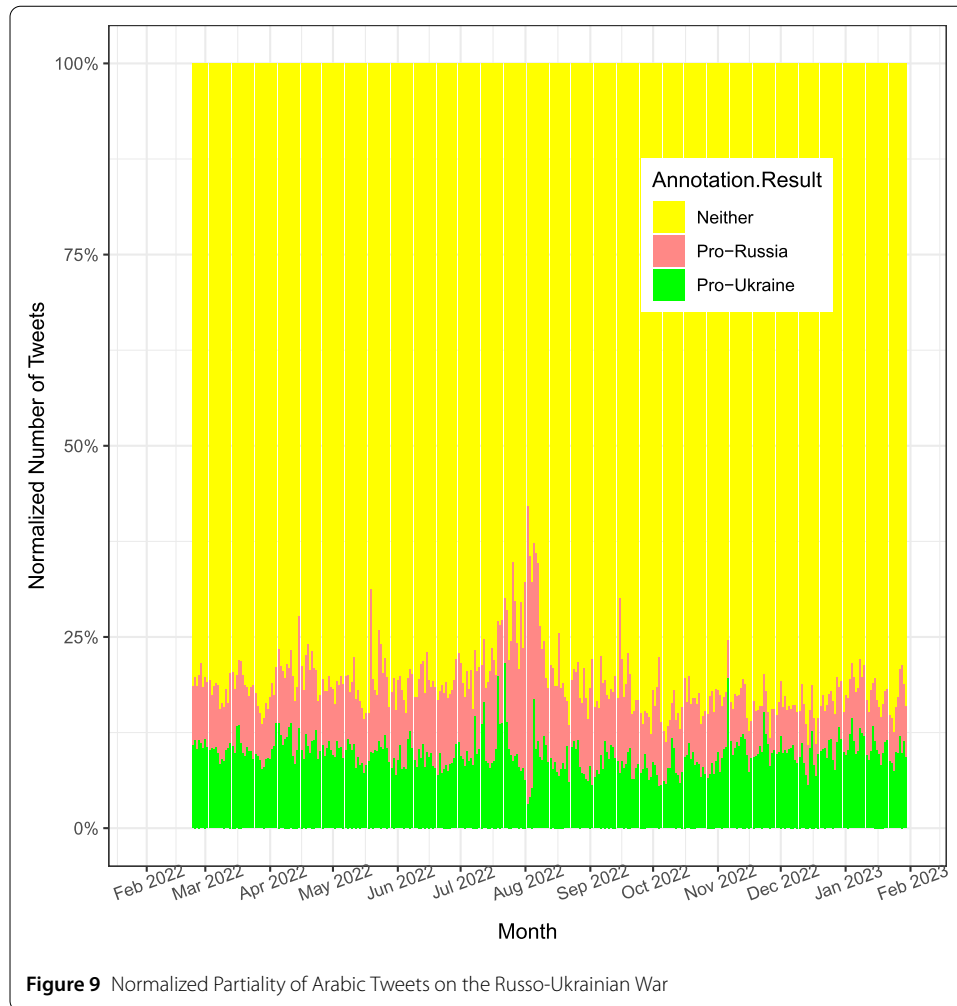
It is apparent that both parties, namely Russia and Ukraine, are almost equal regarding their support (among opinionated people) in the Arab region with a slight shift towards Russia. It seems a bit surprising that the amount of empathy each party gained is almost equal (or even leaning more towards Russia), even though Russia is the aggressor. In order to try to understand that we did finer investigation over the temporal evolution of the partiality/empathy over the course of the war.





Figures 8 and 9 illustrate the evolution of each individual partiality over time since the beginning of the war and its normalized version, respectively. As shown in these figures, the amount of tweets classified as being ‘Pro-Ukraine’ was slightly greater than Pro-Russia’ at the beginning of the war (specifically from Feb 2022 till April 2022) then slowly began to decrease until they nearly converged at the start of June 2022 with a shift happened in the empathy towards Russia in August 2022. Our Interpretation for that is with the initial Russian fierce and surprise attack at the beginning and the amount of refugees who escaped to neighbouring countries, Ukraine gained much empathy. However, by April 2022, Russian intensity has been decreased and with heavy sanctions the U.S. and West have applied on Russia, Russia has begun to gain such empathy with decrease on the Ukrainian side.

There is an upsurge in ‘Pro-Russia’ in August 2022, as Putin, the President of Russia has signed a decree on Thursday 25th of August 2022 in order to increase Russia’s armed forces size from 1.9M to 2.04M [67]. There is a more Pro-Ukraine towards the end of the year (about one year from the start of the war), as Russia has lost a reported 200,000 subjects, including many high-ranking military officials, and Putin was confounded by the successes of the Ukrainian army/citizens [68].



For results validation, a representative dataset of 1000 tweets was selected. Out of these, 780 tweets belong to the testing dataset annotated by the LSTM model, while 220 tweets belong to the training dataset annotated by the lexicon-based approach. The ratio between both datasets complies with the size of the LSTM testing dataset (1.471M) and the lexicon-based training dataset (449K).

Each of the 1000 tweets was manually annotated by everyone of four individuals based on the following criteria:

- Pro-Russia: the tweet either defends “Russia, the Russian army, the Russian president, Russian affiliated organizations/personnel, allies” or strongly criticizes “Ukraine, Ukrainian affiliated organizations/personnel, allies”.
- Pro-Ukraine: the tweet either defends “Ukraine, the Ukrainian army, the Ukrainian president, Ukrainian affiliated organizations/personnel, allies” or strongly criticizes “Russia, Russian affiliated organizations/personnel, allies”.
- Neither: the tweet is not siding with any of the two parties, criticizing both, or not related to the conflict.

Majority voting for the four individuals’ decisions was taken for selecting the final label. The four individuals had the same annotation of the sample labels given by either LSTM model or lexicon-based approach for 70.1% of the cases. The LSTM model was retrained

again with increasing the training dataset size to be about 80% of the whole dataset and the accuracy increased to 77.3%, strongly indicating that increasing the dataset size made the model to learn wider patterns leading to improved performance.

## 5 Comparison with recent works

In this section, we compare our main results and analysis to similar works analyzing Twitter datasets of the Russo-Ukrainian war. Specifically, two seminal works are done in this area; the one authored by Shevtsov et al. [5] and the other authored by Haq et al. [6].

The authors in [5] performed an initial analysis for the number of tweets and users and the corresponding sentiments. This work contains multiple languages; e.g., English, French, German, Italian, Spanish, Japanese, etc. The number of tweets written in Arabic (from 23rd Feb 2022 till 21st Dec 2022) is 411,830 out of 109,785,023 (about 0.38%)<sup>3</sup> while our dataset contains 3,167,210 Arabic tweets (from 23rd Feb 2022 till 31st Jan 2023). Their results showed more positive sentiment towards the Ukrainian side compared to the Russian side during the whole period.<sup>4</sup> The sentiment analysis was performed using a multi-language pre-trained XLM-RoBERTa model.<sup>5</sup>

The authors in [6] presented in the first release, over 1.6 million tweets published during the first week of the crisis (by the 6th of March 2022). The authors did not perform filtration on any language or geo-locations in the data scrapping. Thus, the dataset includes tweets from various regions in different languages worldwide. The daily amount of tweets has an average of about 200K tweets. There are more than 900,000 users in the current version of this dataset. From all of the tweets, more than 1.2 M are retweets. Out of these tweets, 413,254 are unique tweets which were retweeted with mean of 3 retweets for every tweet and standard deviation of 12.04. This paper does not include any results regarding opinion mining, e.g., sentiment analysis, emotion analysis, etc.

## 6 Conclusion

In this work, we have collected a dataset of tweets, in the Arabic language, related to the Russia-Ukraine war and introduced a newly built lexicon to be used as a tool for the given analysis. We did an analysis to understand the reactions of the Arabic-speaking people towards the conflict that has, seemingly suddenly, erupted between Russia and Ukraine. As the events are still unfolding, we study only almost a year of events between 24th Feb 2022 and 31st Jan 2023. To the best of our knowledge, this is the first such work to handle social media response towards this particular conflict in the Arabic language.

The analysis is done through standard tools in text analysis including sentiment analysis which measures the general attitude of the subject being positive/negative/neutral. The second analysis is more elaborate looking into the particular associated emotion(s); where the typical emotions include: Anger, Joy, Fear, Sad, Surprise, and Disgust. The third track of analysis is concerned with partiality analysis, whether the subject favors or is empathetic to either of the two conflicting sides.

The vast majority of tweets' sentiment were negative, which somehow is expected regarding the attitude towards violent events and the tragedies stemming from wars. Regarding emotion analysis, the majority were Null and Mixed Feelings which shows a kind

---

<sup>3</sup><https://alexdrk14.github.io/RussiaUkraineWar/index.html>.

<sup>4</sup><https://alexdrk14.github.io/RussiaUkraineWar/sentiment.html> [5].

<sup>5</sup><https://huggingface.co/vicgalle/xlm-roberta-large-xnli-anli>.

of confusion, especially with the contradicting reports from both sides about the conflict and the uncertainties regarding the unfolding events. The most common emotion is ‘anger’.

We wanted to contribute to the intuitive political understanding of the public opinion towards the two conflicting parties which can be roughly phrased as a confrontation between the east and the west. We did a preliminary investigation using partiality analysis: the bias and/or the empathy of the people towards either of the conflicting parties. We found that the collected tweets somehow were not oriented towards a specific side, with a little bit of lean towards Russia, especially towards the latter phases of the war.

For future work, we aim to collect more data in order to cover a larger temporal span of the conflict, and hopefully to cover the whole wishing it to end very soon. We also would like to dive deeper and do more thorough analysis of emotions and partiality. We want to tackle these two analyses using a machine learning approach with smarter methods for data collection and annotation. We also plan to incorporate media and political theories to have a deeper and more thorough understanding of the historical context and the evolution of the conflict over time. We would like to analyze other sources of textual media such as Facebook posts and blogging articles, as well as other media types such as audio and video.

#### Acknowledgements

Not applicable.

#### Funding

Open Access funding provided by The Science, Technology & Innovation Funding Authority (STDF) in cooperation with The Egyptian Knowledge Bank (EKB).

#### Availability of data and materials

The datasets generated will be available from the corresponding author upon reasonable request.

#### Declarations

##### Competing interests

The authors declare that they have no competing interests.

##### Author contributions

WG formulated the problem of analyzing the Arab reactions towards the Russo-Ukrainian war. MT, MK, AY, SK, and AA performed the data processing and code implementation. All authors analyzed and interpreted the results. All authors read and approved the final manuscript.

##### Author details

<sup>1</sup>Egypt Japan University of Science and Technology, P.O. Box 179, 21934, New Borg El-Arab City, Alexandria, Egypt. <sup>2</sup>Ejada Systems Ltd., 620-622 El Horeya Street, Zizinia, Alexandria, Egypt. <sup>3</sup>Faculty of Engineering, Alexandria University, P.O. Box 21544, Alexandria, Egypt.

Received: 24 April 2023 Accepted: 25 August 2023 Published online: 15 September 2023

#### References

1. Dong X, Lian Y (2021) A review of social media-based public opinion analyses: challenges and recommendations. *Technol Soc* 67:101724
2. Dolan P, Kavetsos G, Krekel C, Mavridis D, Metcalfe R, Senik C, Szymanski S, Ziebarth NR (2016) The host with the most? The effects of the Olympic Games on happiness. *SSRN Electron J*, 1–50
3. Goodwin M, Hix S, Pickup M (2020) For and against brexit: a survey experiment of the impact of campaign effects on public attitudes toward EU membership. *Br J Polit Sci* 50(2):481–495
4. Xue J, Chen J, Hu R, Chen C, Zheng C, Su Y, Zhu T et al (2020) Twitter discussions and emotions about the COVID-19 pandemic: machine learning approach. *J Med Internet Res* 22(11):20550. <https://doi.org/10.2196/preprints.20550>
5. Shevtsov A, Tzagkarakis C, Antonakaki D, Pratikakis P, Ioannidis S (2022) Twitter dataset on the Russo-Ukrainian war. arXiv preprint. [arXiv:2204.08530](https://arxiv.org/abs/2204.08530)
6. Haq E-U, Tyson G, Lee L-H, Braud T, Hui P (2022) Twitter dataset for 2022 Russo-Ukrainian crisis. arXiv preprint. [arXiv:2203.02955](https://arxiv.org/abs/2203.02955)

7. Duho KCT, Abankwah SA, Agbozo DA, Yonmearu G, Aryee BNA, Akomanin O (2022) Exploring the Russo-Ukrainian crisis and its impact on african countries: a cross-regional analysis. *SSRN Electron J*, 1–54. <https://doi.org/10.2139/ssrn.4085903>
8. Al Farhoud YT (2018) The use of Twitter as a tool to predict opinion leaders that influence public opinion: case study of the 2016 United State presidential election. In: Knowledge discovery and data design innovation: proceedings of the international conference on knowledge management (ICKM 2017). World Scientific, Singapore, pp 191–206
9. Karamouzas D, Mademlis I, Pitas I (2022) Public opinion monitoring through collective semantic analysis of tweets. *Soc Netw Anal Min* 12(1):1–21. <https://doi.org/10.1007/s13278-022-00922-8>
10. Adwan O, Al-Tawil M, Huneiti A, Shahin R, Zayed AA, Al-Dibsi R (2020) Twitter sentiment analysis approaches: a survey. *Int J Emerg Technol Learn* 15(15):79–93. <https://doi.org/10.3991/ijet.v15i15.14467>
11. Abu Farha I, Magdy W (2019) Mazajak: an online Arabic sentiment analyser. In: Proceedings of the fourth Arabic natural language processing workshop. Assoc. Comput. Linguistics, Florence, pp 192–198. <https://doi.org/10.18653/v1/W19-4621>. <https://aclanthology.org/W19-4621>
12. LeCun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. *Proc IEEE* 86(11):2278–2324
13. Schmidhuber J, Hochreiter S et al (1997) Long short-term memory. *Neural Comput* 9(8):1735–1780
14. Antoun W, Baly F, Hajj H (2020) AraBERT: transformer-based model for Arabic language understanding. In: Proceedings of the 4th workshop on open-source Arabic corpora and processing tools, with a shared task on offensive language detection, Marseille, France. European language resource association, pp 9–15. <https://aclanthology.org/2020.osact-1.2>
15. Kenton JDM-WC, Toutanova LK (2019) BERT: pre-training of deep bidirectional transformers for language understanding. In: Proceedings of NAACL-HLT. Assoc. Comput. Linguistics, Florence, pp 4171–4186
16. Obeid O, Zalmout N, Khalifa S, Taji D, Oudah M, Alhafni B, Inoue G, Eryani F, Erdmann A, Habash N (2020) CAMEL tools: an open source Python toolkit for Arabic natural language processing. In: Proceedings of the 12th language resources and evaluation conference. European language resources association., pp 7022–7032. <https://aclanthology.org/2020.lrec-1.868>
17. Pasha A, Al-Badrashiny M, Diab M, El Kholy A, Eskander R, Habash N, Pooleery M, Rambow O, Roth R (2014) Madamira: a fast, comprehensive tool for morphological analysis and disambiguation of Arabic. In: Proceedings of the ninth international conference on language resources and evaluation (LREC'14). Springer, Berlin, pp 1094–1101
18. Abdelali A, Darwish K, Durrani N, Mubarak H (2016) Farasa: a fast and furious segmenter for Arabic. In: Proceedings of the 2016 conference of the North American chapter of the Association for Computational Linguistics: demonstrations. Assoc. Comput. Linguistics, Florence, pp 11–16. <https://doi.org/10.18653/v1/N16-3003>
19. Manning CD, Surdeanu M, Bauer J, Finkel JR, Bethard S, McClosky D (2014) The Stanford CoreNLP natural language processing toolkit. In: Proceedings of 52nd annual meeting of the Association for Computational Linguistics: system demonstrations. Assoc. Comput. Linguistics, Florence, pp 55–60
20. Elmadany A, Mubarak H, Magdy W (2018) ArSAS: an Arabic speech-act and sentiment corpus of tweets. In: Al-Khalifa H, Magdy W, Darwish K, Elsayed T (eds) Proceedings of the LREC 2018 workshop "The 3rd workshop on open-source Arabic corpora and processing tools (OSACT)". European Language Resources Association (ELRA), pp 20–25
21. El-Khair IA (2016) 1.5 billion words arabic corpus. *arXiv preprint*. [arXiv:1611.04033](https://arxiv.org/abs/1611.04033)
22. Zeroual I, Goldhahn D, Eckart T, Lakhouaja A (2019) OSIAN: open source international Arabic news corpus-preparation and integration into the CLARIN-infrastructure. In: Proceedings of the fourth Arabic natural language processing workshop. Assoc. Comput. Linguistics, Florence, pp 175–182
23. Habash N, Eskander R, Hawwari A (2012) A morphological analyzer for Egyptian Arabic. In: Proceedings of the twelfth meeting of the special interest group on computational morphology and phonology (SIGMORPHON). Assoc. Comput. Linguistics, Florence, pp 1–9
24. Obeid O, Salameh M, Bouamor H, Habash N (2019) ADIDA: automatic dialect identification for Arabic. In: Proceedings of the 2019 conference of the North American chapter of the Association for Computational Linguistics (demonstrations). Assoc. Comput. Linguistics, Florence, pp 6–11. <https://doi.org/10.18653/v1/N19-4002>. <https://aclanthology.org/N19-4002>
25. Shaalan K, Siddiqui S, Alkhatib M, Abdel Monem A (2019) Challenges in Arabic natural language processing. In: Computational linguistics, speech and image processing for Arabic language. World Scientific, Singapore, pp 59–83. [https://doi.org/10.1142/9789813229396\\_0003](https://doi.org/10.1142/9789813229396_0003)
26. Gabel S, Reichert L, Reuter C (2020) Discussing conflict in social media: the use of Twitter in the Jammu and Kashmir conflict. *Media, War Confl* 15(4):1750635220970997. <https://doi.org/10.1177/1750635220970997>
27. Ouyang Y, Waterman RW (2020) Trump tweets: how often and on what topics. In: Trump, Twitter, and the American democracy. Springer, Berlin, pp 53–87. [https://doi.org/10.1007/978-3-030-44242-2\\_3](https://doi.org/10.1007/978-3-030-44242-2_3)
28. Yum S (2020) Mining Twitter data to understand the human sentiment on hurricane Florence. *J Disaster Emerg Res* 3(2):74–86. <https://doi.org/10.18502/jder.4069>
29. França T, Gomes J, Oliveira J (2017) A Twitter opinion mining gold standard for Brazilian uprising in 2013. In: 32th Simpósio Brasileiro de Banco de Dados (SBBDD). Sociedade Brasileira de Computação – SBC, pp 182–192
30. Magdy W, Darwish K, Weber I (2015) #FailedRevolutions: using Twitter to study the antecedents of ISIS support. *First Monday* 21. <https://doi.org/10.5210/fm.v21i2.6372>
31. Magdy W, Darwish K, Abokhodair N (2015) Quantifying public response towards Islam on Twitter after Paris attacks. *arXiv preprint*. [arXiv:1512.04570](https://arxiv.org/abs/1512.04570)
32. Ren R (2022) Emotion analysis of cross-media writing text in the context of big data. *Front Psychol* 13
33. Gomaa W, Elbasiony R (2020) World perception of the latest events in Egypt based on sentiment analysis of the Guardian's related articles. In: The 4th international conference on advanced machine learning technologies and applications (AMLT 2019). Springer, Berlin, pp 908–917
34. Khalil EAH, El Houbay EM, Mohamed HK (2021) Deep learning for emotion analysis in Arabic tweets. *J Big Data* 8(1):1–15
35. Alqahtani G, Alothaim A (2022) Emotion analysis of Arabic tweets: language models and available resources. *Front Artif Intell* 5

36. Baali M, Ghneim N (2019) Emotion analysis of Arabic tweets using deep learning approach. *J Big Data* 6(1):1–12
37. Meo R, Sulis E (2017) Processing affect in social media: a comparison of methods to distinguish emotions in tweets. *ACM Trans Internet Technol* 17(1):1–25
38. Kang X, Ren F (2016) Understanding blog author's emotions with hierarchical Bayesian models. In: 2016 IEEE 13th international conference on networking, sensing, and control (ICNSC). IEEE, Los Alamitos, pp 1–6
39. Hasan M, Agu E, Rundensteiner E (2014) Using hashtags as labels for supervised learning of emotions in Twitter messages. In: *ACM SIGKDD workshop on health informatics*, vol 34. ACM, New York, p 100
40. Shaaban Y, Korashy H, Medhat W (2021) Emotion detection using deep learning. In: 2021 16th International Conference on Computer Engineering and Systems (ICCES). IEEE, Los Alamitos, pp 1–10
41. Shaaban Y, Korashy H, Medhat W (2022) Arabic emotion cause extraction using deep learning. *Egypt J Lang Eng* 9(2):23–39
42. Gupta N, Agrawal R (2020) Application and techniques of opinion mining. In: *Hybrid computational intelligence*. Elsevier, Amsterdam, pp 1–23
43. Kannan S, Karuppusamy S, Nedunchezian A, Venkateshan P, Wang P, Bojja N, Kejariwal A (2016) Big data analytics for social media. In: Buyya R, Calheiros RN, Dastjerdi AV (eds) *Big data*. Kaufmann, Los Altos, pp 63–94. <https://doi.org/10.1016/B978-0-12-805394-2.00003-9>. <https://www.sciencedirect.com/science/article/pii/B9780128053942000039>
44. Denecke K (2008) Using SentiWordNet for multilingual sentiment analysis. In: 2008 IEEE 24th international conference on data engineering workshop. IEEE, Los Alamitos, pp 507–512
45. Hamouda A, Rohaim M (2011) Reviews classification using SentiWordNet lexicon. *Online J Comput Sci Inf Technol* 2(1):120–123
46. Chalothorn T, Ellman J (2012) Using SentiWordNet and sentiment analysis for detecting radical content on web forums. *Nrl Northumbria Ac Uk* 1
47. Ohana B, Tierney B (2009) Sentiment classification of reviews using SentiWordNet. *Proc IT&T* 8
48. Badaro G, Baly R, Hajj H, Habash N, El-Hajj W (2014) A large scale Arabic sentiment lexicon for Arabic opinion mining. In: *Proceedings of the EMNLP 2014 workshop on Arabic Natural Language Processing (ANLP)*. Assoc. Comput. Linguistics, Florence, pp 165–173. <https://doi.org/10.3115/v1/W14-3623>
49. El Abbadi N, Khedhair A, Al-Nasrawi A (2011) Build electronic Arabic lexicon. *Int Arab J Inf Technol* 8:137–140
50. Youssef M, El-Beltagy SR (2018) MoArLex: an Arabic sentiment lexicon built through automatic lexicon expansion. *Proc Comput Sci* 142:94–103. <https://doi.org/10.1016/j.procs.2018.10.464>
51. El-Beltagy SR (2019) WeightedNileULex: a scored Arabic sentiment lexicon for improved sentiment analysis. In: *Computational linguistics, speech and image processing for Arabic language*. World Scientific, Singapore, pp 169–186. [https://doi.org/10.1142/9789813229396\\_0008](https://doi.org/10.1142/9789813229396_0008)
52. Soliman AB, Eissa K, El-Beltagy SR (2017) AraVec: a set of Arabic word embedding models for use in Arabic NLP. *Proc Comput Sci* 117:256–265
53. Mohammad S, Bravo-Marquez F, Salameh M, Kiritchenko S (2018) Semeval-2018 task 1: affect in tweets. In: *Proceedings of the 12th international workshop on semantic evaluations (SemEval-2018)*. Assoc. Comput. Linguistics, Florence, pp 1–17
54. Quan C, Ren F (2010) A blog emotion corpus for emotional expression analysis in Chinese. *Comput Speech Lang* 24(4):726–749
55. Scherer KR, Wallbott HG (1997) The ISEAR questionnaire and codebook
56. Gomaa W (2021) Analysis of Arabic songs: abdel ElHalim as a case study. In: *Advanced machine learning technologies and applications: proceedings of AMLTA 2021*. Springer, Berlin, pp 385–393
57. Gomaa W (2023) Lyrics analysis of the arab Singer Abdel ElHalim Hafez. *ACM Trans Asian Low-Resour Lang Inf Process* 22(2):1–27
58. Tang L, Liu H (2005) Bias analysis in text classification for highly skewed data. In: *Fifth IEEE International Conference on Data Mining (ICDM'05)*. IEEE, Los Alamitos, p 4
59. Recasens M, Danescu-Niculescu-Mizil C, Jurafsky D (2013) Linguistic models for analyzing and detecting biased language. In: *Proceedings of the 51st annual meeting of the Association for Computational Linguistics (volume 1: long papers)*. Assoc. Comput. Linguistics, Florence, pp 1650–1659
60. Demata M (2018) Manipulation and partiality in Italian translations of foreign news about Italy: three case studies. *ESP Across Cult Open Access Mag* 15:27–39. Edipuglia
61. Mahmood A, Khan HU, Ramzan M (2020) On modelling for bias-aware sentiment analysis and its impact in Twitter. *J Web Eng* 19:1–28
62. Yang K-C, Hui P-M, Menczer F (2022) How Twitter data sampling biases US voter behavior characterizations. *PeerJ Comput Sci* 8:1025
63. Gross TW (2020) Sentiment analysis and emotion recognition: evolving the paradigm of communication within data classification. *Appl Mark Anal* 6(1):22–36. <https://doi.org/10.6084/m9.figshare.14095603>
64. Saad MK (2015) Mining documents and sentiments in cross-lingual context. PhD thesis, Université de Lorraine
65. Strapparava C, Mihalcea R (2007) Semeval-2007 task 14: affective text. In: *Proceedings of the fourth international workshop on semantic evaluations (SemEval-2007)*. Assoc. Comput. Linguistics, Florence, pp 70–74
66. Devlin J, Chang M-W, Lee K, Toutanova K (2018) BERT: pre-training of deep bidirectional transformers for language understanding. *arXiv preprint. arXiv:1810.04805*
67. Reuters – World/Europe. Putin signs decree to increase size of Russian armed forces. <https://www.reuters.com/world/europe/putin-signs-decree-increase-size-russian-armed-forces-2022-08-25/>
68. Harvard News. How does Ukraine war end. <https://news.harvard.edu/gazette/story/2023/02/how-does-ukraine-war-end-experts-say-2023-could-prove-decisive-dangerous/>

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.